(5) 169 200

63-4-4

(6) GOAL ORIENTED MODELS

FOR ACCOUNTING AND CONTROL

by

Yuji Ijiri

Carnegie Institute of Technology

Pittsburgh 13, Pennsylvania

Scale 5

GRADUATE SCHOOL of INDUSTRIAL ADMINISTRATION

William Larimer Mellon, Founder

$21.00

(4) $21.00 (5) 169 200

(18) O.N.R. (19) Research Memorandum No. 114

(7) NA
(8) NA
(9) NA
(12) IV.
(13) NA
(14) NA
(17) NA
(20) U
(21) Doctoral thesis

(6) GOAL ORIENTED MODELS

FOR ACCOUNTING AND CONTROL,

(10) by

Yuji Ijiri

ls, C.

(11) May, 1963

Carnegie Institute of Technology

Graduate School of Industrial Administration

This paper was written as part of the contract "Planning and Control of Industrial Operations," with the Office of Naval Research and the Bureau of Ships, at the Graduate School of Industrial Administration, Carnegie Institute of Technology. Reproduction of this paper in whole or in part is permitted for any purpose of the United States Government. (15) Contract Nonr-760 (01), (16) Project NR-047011.

GOAL ORIENTED MODELS
FOR ACCOUNTING AND CONTROL

A Dissertation

Presented to

the Faculty of

the Graduate School of Industrial Administration

Carnegie Institute of Technology

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Yuji Ijiri

May 1963

ACKNOWLEDGEMENT

Grateful acknowledgment is made of the encouragement and criticisms by the members of the thesis committee, Professors W. W. Cooper, chairman, N. C. Churchill, K. J. Cohen, and H. A. Simon of the Graduate School of Industrial Administration, Carnegie Institute of Technology. In particular, my greatest debt is to Professor Cooper for his numerous insightful comments and suggestions as well as for the time he unstintingly gave me to help my thinking on accounting, managerial, and mathematical matters. Perhaps even more important, however, was the lesson he taught me, by examples, of the necessity of searching for constant improvement in my thinking and ways of presenting my ideas as an integral part of the process of research itself.

Professors A. Ben-Israel and G. L. Thompson helped me with careful readings and suggestions on the mathematical development presented in the appendices of this thesis. In particular, Professor Ben-Israel, who is a coauthor of the paper reproduced in Appendix B of the thesis, helped me in working out the theories of generalized inverse.

Thanks are also due to Professor A. Charnes and my colleagues, F. K. Levy and L. C. Lyon, who are coauthors of the papers which have been incorporated in a part of the thesis, as well as to Dean R. M. Cyert of the Graduate School of Industrial Administration for his interest in my endeavors while at the School.

This study has been supported partly by the Ford Foundation Predoctoral Fellowship and partly by the Office of Naval Research, and their support is herewith gratefully acknowledged.

Mrs. Grace Couchman and Miss Barbara Barkovich helped in typing a part of the final manuscript and thereby helped me through a period of crisis when completing the thesis.

I also wish to take this opportunity to thank Professor Taminosuke Nishimura of Doshisha University, Kyoto, Japan, who initially guided me to the accounting profession and who has continously advised me to look into "Lebens goldner Baum" behind the "grau Theorie" of accounting.

Finally, credit is also due to my wife, Tomo, who, in addition to typing most of the preliminary and final manuscripts, has shared the strain of the thesis work willingly and without any complaint.

ABSTRACT

GOAL ORIENTED MODELS

FOR ACCOUNTING AND CONTROL

Yuji Ijiri

1. The Main Problems addressed by the Thesis:

This thesis is directed toward analyzing fundamental factors in accounting for control without being restricted to the conventional monetary accounting systems, and constructing mathematical models by which planning processes and control accounting processes are tied in together for better coordination of the two processes.[1]

2. Accounting for Planning and Accounting for Control:

The classification of accounting into "accounting for planning" and "accounting for control" is first discussed. The basic requirement of the former is to preserve maximum flexibility of data at minimum cost whereas that of the latter is to maintain uniformity and consistency of data. The thesis mainly deals with the latter, even though the outcome from the models in the thesis can, of course, be used for planning purposes as well as control purposes of management.

[1]The main results of the study are listed in Section 1 of Chapter VII in the thesis with indices of chapter and section numbers in which the results are discussed.

3. Planning Processes:

Planning is viewed as a process of translating or decomposing management "goals" into "subgoals" which are more operational and controllable by management and their subordinates than goals. For example, the role of breakeven analysis is to translate the management goal of profit attainment into a subgoal (sales volume) which is more operational and controllable than profit since sales volumes do not depend upon cost factors.[2] Such goals and subgoals of management may be monetary or non-monetary, optimizing or satisficing, single or multiple. The analyses and the model construction in the thesis proceed taking into account all these characteristics of management goals and subgoals without restriction to any particular goals or subgoals. Goal programming and generalized inverses of matrices are utilized in the mathematical models in the thesis. Special problems of incompatible multiple goals are discussed together with a way of deriving a solution by ordering and weighting of goals even when incommensurable goals are involved.

4. Control Accounting Processes:

Accounting for control, or control accounting, is viewed as a process of generating feedbacks or indicators as to where a

[2]The extension of the simple breakeven analysis toward a piecewise linear breakeven model together with an analysis of fixed and variable cost components, comparison of a piecewise linear total revenue curve in the above extended breakeven model with conventional continuous and everywhere differentiable total revenue curve, substitution of the concept of elasticity by "marginal revenue head," etc., are also discussed in connection with the breakeven analysis.

management stands relative to its goals. The central problems of control accounting as described in the thesis are: (i) knowing the relationship between subgoals and goals, what relationship between subgoals and indicators is critical in order to determine the degree of goal attainment from knowledge of indicator values and (ii) if the goal level cannot be determined uniquely from an indicator value, how does the reading of an indicator value help to narrow the possible range of the goal level that has been attained? The concept of an "indicator-goal divergence coefficient" was developed relative to the former problem whereas "indicator-goal control chart" was designed to answer the latter problem. Here again, the analyses and the model construction proceed without restriction relative to any particular indicators.

5. The Double-entry Bookkeeping System:

After the above analyses of the fundamental factors of control accounting, the basic principle of the double-entry bookkeeping system is reviewed by means of a spread sheet, an incidence matrix, and an accounting network. In particular, uses of multiple measuring units in a double-entry bookkeeping system are discussed, as when, say, physical as well as cash flows are to be accounted for and qualitative as well as quantitative designations are to be accomodated.

6. Spread Sheet Planning:

Then, data available under the double-entry bookkeeping system are utilized as subgoals into which the management goals are to be translated, and these subgoals are incorporated in the mathematical models of planning developed in earlier chapters. Such planning is called "spread sheet planning" since these subgoals are elements in a spread sheet of the double-entry bookkeeping system. By spread sheet planning it is possible to incorporate control accounting processes in dealing with planning in any level of detail, producing such managerially useful data as projected balance sheet and projected income statements when plans are executed, accounting statements of dual evaluators, indicator-goal control charts, iso-profit price-volume curves, etc. Practical uses of the spread sheet planning were effected through an application to a firm in the Carnegie Management Game which provides enough complexity for a reasonable test. Then, the problems of multiple presiod spread sheet planning are discussed as a way to combine, simultaneously, capital budgeting and operating budgeting which are ordinarily treated separately in the practice of budgeting.

7. Mathematics in the Thesis

Even though mathematics per se is not the main concern of the thesis, it involves some topics which would be of interest from mathematical view point, too.

In the thesis, the generalized inverse of an arbitrary matrix is utilized. When a matrix is singular or rectangular, the ordinary

inverse of the matrix does not exist. However, the generalized inverse exists uniquely for any matrix, singular or non-singular, square or rectangular, thus opening a new and wide area in applying linear mathematical tools. In view of the fact that few applications of generalized inverses have been made even in the sophisticated area of management science, an introductory analysis of generalized inverses was made in Appendix A as a bridge between elementary linear algebra and the theories of generalized inverse. A discussion of methods for computing generalized inverse is contained in Appendix B, together with a flow chart and a GATE computer program for calculating generalized inverses by the "elimination method." Main interest focuses, however, on the use of these inverses for accounting processes since, in general, such processes are associated with matrices that do not have ordinary inverses.

A method of deriving an indicator-goal control chart is developed in Appendix C of the thesis. This can also be used as a method of generating all the extreme points of a convex set mapped into a two dimensional space. Since the method was developed as an extension of a dual evaluator analysis, only minor modifications in the simplex routine are required for the method. The main point, of course, is to secure a way of actually implementing these indicator-goal control charts and related traditional instruments of management control.

8. Conclusions:

The thesis concludes with remarks on relations between traditional accounting and some currently popular proposals for its alteration. Further areas of research and development are also delineated.

TABLE OF CONTENTS

CHAPTER PAGE

APPENDIX PAGE

LIST OF FIGURES

FIGURE		PAGE

LIST OF TABLES

CHAPTER I

INTRODUCTION

1. Motivation Behind the Thesis

This thesis analyzes relationships between the planning processes of an organization and the accounting processes it uses to generate control data. That is, the concern here is with processes of accounting for control. The aim is to derive criteria for better coordination between the two processes--planning and accounting for control.

The basic purpose of accounting for control, or simply control accounting, is to supply data feedbacks so that a firm's management can decide, from time to time, as to where their organization stands relative to its objectives, or goals. This involves such processes as setting standards, measuring performance, comparing actual performance with standards, reporting deviations from standards, etc.

Control accounting is actually only one branch, although a very important one, of accounting defined in the broadest sense. If accounting is to be grossly classified into financial accounting and management accounting, control accounting may be classified with the latter by virtue of the fact that the use of the data derived from control accounting is mostly internal. However, from the standpoint of overall managerial objectives, or goals--such as profit, etc.--

there is a rather close relationship between control- and financial-accounting in that the details of system design and reporting are formulated and judged relative to these total ends and the inner details of the system are rearranged or adjusted accordingly. To this extent the view of accounting for control purposes differs from the ordinary characterizations of managerial accounting insofar as the latter are oriented to the needs of individual decision makers in the separate parts of a total system.

For purpose of this thesis, in fact, a better classification is obtained by distinguishing between "accounting for planning" and "accounting for control." In accounting for planning, the main problem is directed towards the design of systems which supply information on the firm's operations and on its environment (not necessarily in any specific way related to management goals), in the moxt flexible and economical way. The measure of economy is relative to minimum time or minimum cost and the measure of flexibility is relative to the systems ability to supply the varying needs of management for data in terms of any planning problems that may be relevant. In a make-or-buy decision, for example, management may want to have data such as the opportunity costs of producing or buying the products, the capacity of the currently available machines, capacity of the suppliers, availability of labor, sales forecasts, storage limitations, and so on. Of course, all of this information may not be available in the books of account or the statistical records

of the accounting department, but a wholly flexible system would nevertheless have it all identified and available as required. Of course, cost needs to be considered, too. Hence the main issue for accountants concerning such information requirements from management is either how to obtain maximum data flexibility at a given cost or, vice versa, to minimize the cost of attaining a given degree of flexibility so that a wide variety of information (e.g., average or expected cost, maximum and minimum costs, total cost, etc.) may be prepared in either case from prime information (e.g., costs on individual jobs) with minimum cost or time. On the other hand, it is hard to predict what kinds of information management may need in planning processes that are not repetitive. This adds a further dimension of uncertainty which admittedly clouds and complicates the issues with respect to system design for planning in terms of the previous criteria and so it is difficult to construct realistic general theories and models for system design in accounting for planning.

It is not the object of this thesis to deal with these problems in accounting for planning. The purpose of the preceding discussion is rather simply to provide a general sketch of these problems and then to contrast them with problems in accounting for control. Contrary to the basic demand of flexibility in accounting for planning, the basic demand in accounting for control is uniformity and consistency. This will subsequently be developed in more detail but, for the moment, we may say that this demand arises from the related needs of "measurement"

and "comparison" in a complex of activities over continuing periods of time. We may then use this as a way of explaining the tendencies in accounting which have been directed towards uniformity and consistency in practice as well as in the accounting theories, especially in financial accounting, which have been developed systematically.

One purpose of control accounting is to measure and report "actual" performance of management and their subordinates relative to overall and subsidiary management goals. It follows that we must know what these goals are before we can set up a control accounting system which generates feedbacks for the processes of goal attainment.

From the preceding discussion we can again see that there are close relations, at least on a functional level, between control and financial accounting. In fact, in the sense just discussed, financial accounting may be viewed as control accounting for cases in which the management goal is to attain a maximum or satisfactory level of profit. Hence, from this standpoint, standard accounting theories on asset valuation, income determination, etc., may be viewed as directed toward establishing a control system which generate feedbacks in terms of special goals such as the attainment of satisfactory level of "profit." This control may be viewed as operating on management via its stockholder reports and other related devices. It may also be viewed as a social and governmental control and doubtless has other dimensions as well. Insofar as this is true we can then

begin to understand some of the pressures towards uniformity and consistency, when financial accounting is viewed from a control standpoint, and then we can also begin to understand why this may sometimes be purchased at the cost of some decrease in flexibility from a purely management standpoint.

If we now view top management as being laced to its external environment by controls such as these then we can also begin to obtain a clearer assessment of accounting control within the managerial hierarchy itself. For instance, profit is only one of numerous possible goals which management may have. This may be true even if profit is regarded as an ultimate control on this management because in most firms it will be necessary to consider also the goals of other tiers of management as well. The latter, of course, may not be immediately and operationally translated into overall total profit. For example, a sales manager's goal may be to attain a given target sales volume assigned by the top management, or a production manager's goal may be to fulfill the production requirements with minimum costs. Note that control accounting exists not only for top management but also for their subordinates at various vertical and horizontal stages of an organizational hierarchy. This means that the concept of "entity" in financial accounting should not be restricted to the firm in its entirety when we are talking about control accounting in general.

From a formal theoretical standpoint it may be true that there is no barrier to calculating the profit of a firm and using the

resulting imputations to different individuals instead of calculating say, a figure on sales volume for a sales department and using this as a control. In practice, however, accountants generally prepare control feedbacks for various types of goals for various segments of an organization and managers rely on these feedbacks and adjust their operations accordingly. Sometimes these reports are even used as a basis for promotion within or across departments and penalties as well as rewards are assessed against different individuals on the basis of the "feedbacks" prepared by accountants. It would seem that some attempt at theorizing or explanation should be essayed here but virtually all of the main theories in accounting are directed solely towards the goal of profit for a firm as a whole. This is true of the asset and equity accounts as well as the income statement [1/] reports and even of the cost and "managerial

1/ See especially American Institute of Accountants, 1952, as well as the more recent AICPA report by R. T. Sprouse and Maurice Moonitz (1962) which would axiomatize all of accounting by the basis of asset valuations in terms of income flows.

accounts" as well. But accounting for control is also important and a completely adequate theory of accounting should accommodate it as well. From this standpoint at least, it is our belief that accounting theories of asset valuation, income determination, etc., should be developed and restated as only one part of a new and higher dimension of accounting theory in which accounting for control, and perhaps

other topics, too, will find a place. It is our belief that when this is done it will be found necessary to relax and alter ideas that have now become relatively fixed on such things as monetary valuation conventions, etc., in order to broaden the perspective of accounting theorists beyond their limited areas of current attention.

2. Lines of Attack

As has already been observed, there is very little formal theory available in the area of accounting for control.[1/] This thesis is intended as a contribution in this direction. It considers accounting for control, per se, as a prime topic of interest and attempts to formalize aspects of the problems encountered for control data. This is done by means of a general treatment that is not restricted to any particular type of management goal. Since planning processes and control processes are so closely related the first part of the thesis is concerned with an analysis of planning processes. These are then related to control processes after which the latter is then made the center of subsequent discussion.

From the standpoint of this thesis, planning is considered as a process of decomposing given management goals into a set of subgoals (or means) which are more operational and controllable by management or their subordinates than goals. Here, for example, budgeting is viewed as a process of constructing a set of subgoals for a given goal of profit attainment. Such subgoal-goal notions or means-ends notions are, of course, only relative. Each subgoal thus derived may be regarded as a goal given to a subordinate, on which the subordinate

---

1/ Notable attempts have recently been made, however, in the cases of A. Stedry (1960) and C. P. Bonini (1962). An earlier, unpublished, attempt may also be found in the study by W. W. Cooper, (1950).

constructs sub-goals which will achieve the given subgoal.[1] In such

[1] Refer to March and Simon, 1958.

a string of subgoal-goal relationships, we then take one segment as representative and consider the planning process as one which is directed towards deriving a set of subgoals that will collectively achieve the given goals. For example, the given level of profit (a goal) may be derived by a combination of sales volumes of several kinds of products (subgoals). Such a set of subgoals which collectively attain the given goals need not be unique, however, and when this is the case there is opened, as a consequence, some discretion and problems of further choice by management and accountants.

Of course a derivation of the relationships between subgoals and goals must proceed by reference to technological, economic, legal, or social environmental considerations if it is to be really relevant. In any event, then, the central problem of control accounting becomes: how we should measure performance in the subgoals in order to determine performance in terms of a given goal. For example, measuring the sales volume of each of several products that contribute to a goal may be used if it can be related in some way to other measures that are available for determining the goal level which has been achieved. Notice, however, that instead of measuring sales volume of each one of several products, the total dollar

sales volume of all the products may be a perfectly good feedback from which the profit level may be inferred, or the total weight of products shipped may very well be a good indicator. The problem, here, is then: knowing the relationship between subgoals and goals what relationship between subgoals and feedbacks, or indicators, is critical in order to determine the goal level from knowledge of indicator values?

The latter topic will subsequently be amplified in considerable detail as a main topic of this thesis. Here, however, we should note that several other topics are also treated for their bearing on a better coordination between accounting and management. Numerous aspects of the latter topic could be treated including new methods of record keeping and reporting. Here, however, we have restricted ourselves, in the main, towards innovations that will help to produce more flexible uses of the kinds of data that are likely to be available from the accounting practices now in use. This treatment has not been restricted to the area of control accounting. Neither has it been restricted to the confines of the present mechanics of accounting for planning purposes. Thus we suggest alternative ways of "booking" and "reporting" accounting transactions even when this departs from practices that are now in wide use. We also apply "accounting mathematical techniques" such as linear programming, etc. The main orientation, however, is towards a use of these techniques for securing increased flexibility in managerial uses of current accounting

data. Thus, for instance, we utilize standard kinds of accounting cost categories--variable and fixed, semi-variable and semi-fixed, etc.--along with related projections of "historical" (or projected "outlay") costs to implement the mathematical formulations directly. Then we rely on known mathematical properties of these "models" to obtain the related opportunity costs in readily ascertainable form.

3. Mathematics Used in the Thesis

At this point we want to mention some of the other motivations that have led to the kinds of mathematics we shall use in our analyses. As mentioned earlier , we are interested in relationships among goals, subgoals, and indicators in general forms without specifying what actually they are. This means that we are not concerned with particular quantitative details of these goals, subgoals, and indicators, but rather with their qualitative relationships. For concreteness we shall refer to particular goals such as profit, sales volume, or total costs, etc. But we do not mean thereby to restrict ourselves to these particular illustrations. Moreover, we also do not want to restrict ourselves to specific dimensions like total dollar sales, or the total weight of products shipped, number of sales orders, etc., and in fact we do not even want to restrict ourselves to the dimensions of ordinary numbers. For instance, we may, like in thermometry analysis, be concerned only with directions or rates of change in a measurement of qualitative properties like "temperature" and we are not always concerned with whether different readings that are obtained at different times can be added or subtracted in any meaningful manner. We are, however, concerned with issues like calibration in the sense that we want our scale to be uniform and consistent for the readings that will be obtained by different observers, say, on a replicated application in any particular situation.

For effecting analyses at this level of abstraction it is both

convenient and efficient to use the language of mathematics. But, of course, this should be susceptible of verbal explanation, too, in order that the content of the models may be understood and assessed. Therefore, in our analyses we first construct some mathematical models that will reflect the planning and accounting processes that can be developed our analyses of these models. But the results of these explorations are then always translated into management terms by reference to special examples.

The mathematical prerequisite required to read this thesis are elementary linear algebra and an introductory knowledge on linear programming--e.g. at the level contained in Chapters IV and V of Charnes and Cooper (1961), or Chapters IV and V of Kemeny, Snell, Mirkil, and Thompson (1959). However, Chapter VI of this thesis can be read without a knowledge of linear algebra or linear programming--even though a linear programming technique is there used to derive solutions to a planning problem using accounting data.

The mathematical models developed in the thesis are all linear models. This does not mean that actual planning and accounting processes are all linear. However, linear models seem to have two essential properties: they are satisfactory as a first approximation to the actual processes, and yet they are analytically simple enough to produce interesting theoretical results. There are, moreover, a wide variety of nonlinear <u>problems</u> that have proved amenable to treatment by only linear <u>means</u>, and it seems to unwise to abandon

possibilities from this quarter ab initio. Computational convenience, ease of interpreting the results obtained from the models in terms of management language, ease of making sensitivity analyses, and so on, are also advantages that can be secured by dealing only with linear models.[1]

1/ See the discussions on justification of the use of a piecewise linear model in Section 2-(3) of Chapter II in the thesis.

For this reason, we have concentrated our efforts on making the most effective use of the concepts and the theories of linear algebra in our analyses. In particular, we have used the concepts and the theories about generalized inverses of arbitrary matrices. In linear systems, we often encounter situations in which we need to derive the inverse of a matrix. If the matrix is singular or rectangular, however, its ordinary inverse is not defined. In order to get rid of this non-singularity restriction on the matrices that are of interest here, we did not want to restrict ourselves further to completely determined systems. Hence we used the theory of generalized inverses which exist for any matrices, even those that are singular, rectangular, or that contains only zero elements.

In view of the fact that few applications of generalized inverses have been made, even in sophisticated areas of management science, we developed, in Appendix A, an introductory analysis of generalized inverses as a bridge between elementary linear algebra and the theories of generalized inverses, thus avoiding a long

mathematical digression in the main part of the thesis. Even though, in this way, we proceeded in the main part of the thesis assuming the reader's knowledge of properties of generalized inverses, we reproduced some of the properties of generalized inverses in the discussions and footnotes of the main part of the thesis so that those who want to read this thesis without referring to Appendix A may do so without undue difficulty.

A discussion of methods for computing generalized inverses is contained in Section 6 of Appendix A and in Appendix B, together with a flow chart and GATE computer program for calculating generalized inverses by the "elimination method." This was done, in part, because it was hoped that this might be of help for some of the practical applications of these ideas.

4. The Structure of the Thesis

Following this introductory chapter, we commence in Chapter II by reviewing a familiar, extremely simple, linear model form of breakeven analysis. This provides a convenient beginning for much of the subsequent analysis. For instance, as is noted in Chapter II, the breakeven model carries with it the notion of "satisficing" as an objective in contrast with the "optimizing" that is normally assumed in much of formal economic analysis. We also use this opportunity to note that, essentially, the role of the breakeven analysis is to translate a management goal (i.e., a breakeven level of profit) into a more operational subgoal (a sales volume).

In Section 2 of Chapter II, the simple linear breakeven model is extended to a "piecewise linear"--i.e., nonlinear accounting--breakeven model as one approach to a wide variety of non-linear revenue and cost curves. We also review some of the distinctions between fixed, variable, semi-variable and semi-fixed costs in order to see how the standard versions of such cost curves, may be approximated, even when nonlinear, by a set of only "variable cost components." Of course, the approximation may also be "real" especially when, as is often the case, it is not wholly clear that the nonlinear curve is itself only an approximation which was originally developed in order to obtain access to convenient properties like continuity, differentiability, etc. In any event, we also try to justify our

approach via piecewise linear models as possibly superior to the approach by the use of continuous and everywhere differentiable functions which are ordinarily used in formal economic analysis. Of course, this can raise problems with reference to other notions such as "elasticities," etc. But even when this is the case, however, other alternatives may be developed. For instance the lack of a well defined elasticity which occurs at points where a derivative does not exist may be compensated by referring back to the original idea of one economic magnitude being incremented as a result of incremental variations in another economic magnitude and so for this reason we develop other ideas such as the "marginal revenue head" which is discussed in this section of Chapter II.

In Section 3 of Chapter II, we deal with cases in which a given goal can be attained by a combination of two or more subgoal variables. We there introduce the notions of a "special solution" and a "general solution factor" by means of the generalized inverse of a matrix and the null space basis of the transformation associated with the matrix. We then go into the analysis of goals in Section 4 by reference to the special type of linear programming which goes by the name of "goal programming."

Chapter III, in its first section, deals with cases in which environmental constraints are imposed upon subgoal variables in goal programming. There analyses are made on various types of functionals which may be incorporated in a goal programming model together with their managerial implications. In Section 3, the goal programming approach

is extended to analysis of multiple goals, together with possible orderings and weightings of multiple goals when these goals are found to be incompatible. The same problem of multiple goals, including the discussion on ordering and weighting of incompatible multiple goals, is also approached by means of generalized inverses in Section 4. We then note differences that arise between these two approaches to incompatible multiple goals after which we relate them by means of the fact that our version of a goal programming approach minimizes the sum of the absolute value of the deviation from each goal level whereas the generalized inverse approach minimizes the "Euclidean distance" from a set of goal levels. Practical applications of such models of planning processes are suggested by means of a set of examples.

In Chapter IV, after setting up models of planning processes, we discuss the control problems of generating feedbacks for goal attainment. We do so by first setting up a model for accounting processes of generating feedbacks and then we derive some criteria which can be imposed upon accounting processes in order to be able to determine the degree of goal attainment from a set of feedbacks.

These topics occupy the first 3 sections of this chapter. In Section 4, we show that when constraints are imposed upon subgoal variables the requirements on the accounting process may be relaxed. In this case, the actual goal level may not be uniquely determined from the reading of an indicator or feedback prepared from the accounting processes, but the range for the actual goal level (maximum and minimum) can be determined because, for practical purposes knowledge of this range may be satisfactory. A way to derive a chart which shows such a relationship between a goal and an indicator is developed in Appendix C along with a computational algorithm we have devised for this purpose. Managerial uses of such a chart, which is called an "indicator-goal control chart", are discussed. This chart is, we think, particularly informative in comparison with some of the devices--e.g., breakeven charts, profit graphs, etc.--that are now employed. For instance it generally substitutes control ranges for the precise point indicators of the latter charts in such a way that subordinates may be given further leeway on their accomplishments while they are, nevertheless, instructed specifically how their performances will be evaluated within the indicated ranges. We also develop a coefficient, called an "indicator-goal divergence coefficient" (or simply a "divergence coefficient") which shows the degree of divergence between the relationship between subgoals and a goal as it, at the same time, shows the relationship between subgoals and an indicator.

In Chapter V and VI, we try to combine the techniques used in analysis of goals with data that are normally available from the usual double-entry bookkeeping system. As has already been suggested, we hope, in this way, to show how to extract more useful managerial data from the data generated by currently available systems. In preparation for this we review, in section 2 of Chapter V, the basic equation of double-entry bookkeeping.

In particular, we note the fact that multiple measuring units can perfectly be accomodated in the double entry bookkeeping system. Then, in Section 3, we represent the relationship between transactions and asset-equity balances by means of a so-called "spread sheet," which, in turn, is related, as we show, to the mathematical constructs of an incidence matrix and an accounting network. Next, we use as subgoal variables in the analysis of goals the variables set up for each cell in a spread sheet. We refer to these as spread sheet variables and observe that these can be placed in exact correspondence with the branches of a rigorously defined "accounting network." By means of these ideas we are then able to develop an example of planning based on spread sheet variables, which can be called "spread sheet planning." After this has all been done we come to Section 5 of this same Chapter where we briefly review the problem of aggregation of activities in connection with "spread sheet planning". This is necessitated because "spread sheet planning" will generally involve some need for basing planning decisions on aggregates in which spread sheet variables are used as aggregates of individual transactions.

Chapter VI is a report on an experimental application of the "spread sheet planning" ideas that are discussed in Chapter V. In view of the fact that many of these ideas were still only in an experimental stage, and were not even wholly and explicitly formulated at the time of this study, it did not seem prudent to effect an immediate contact with the "real" world of a going business in its full complexity. Hence, as a test for the ideas of "spread sheet planning" which also allowed further room for experimentation, this application was effected by reference to the operation of a firm in the so-called Carnegie Tech Management Game. This is a game which possesses sufficient complexity so that it did provide a challenge while it allowed us, at the same time, to trace an array of "typical" business features such as production, marketing, finance, etc., in an experimental context. In Section 4 and 5, of this chapter, after discussion of the model construction and the results obtained from the model, we show, by means of a dual evaluator as well as a sensitivity analysis, how further flexibility may be readily added to the data streams produced from the usual double-entry bookkeeping system when the latter are systematized and exploited by suitable models with such mathematical aids. The model dealt with in Sections 1-5 of Chapter VI is only a single period model. It was not deemed desirable to extend this in all detail to a corresponding multiple-period model. Instead, the implications of extending the

model to a multiple period version are discussed in Section 6 of Chapter VI in order to note the fruitful results which can thereby be obtained. This also carried with it some further implications with respect to current practices. For instance, such a multiple period model via spread sheet planning suggests that capital budgets and operating budgets should be combined for simultaneous consideration in one model. But, then, it should also be noted that this may be done without requiring drastic changes to be made in the more fundamental principles and practices of double-entry accounting as it is currently practiced. That is, the distinction here is one of procedure rather than substance and so we again avoid any need for distinguishing between the requirements of control and other aspects of accounting such as the planning decisions covered in this chapter.

In Chapter VII, all of the main findings of our studies are summarized, together with our conjectures as to where these results may be used to guide some of the future developments that should be considered for accounting in the future.

CHAPTER II

ANALYSIS OF GOALS: AN EXTENSION OF BREAKEVEN ANALYSIS

1. Introduction

The purpose of this chapter is to review some of the basic ideas underlying simple breakeven analysis by a profit graph, and then to extend the analysis by means of the so-called "generalized inverse" of an arbitrary matrix.

The simple profit breakeven analysis can be developed as follows: Suppose a firm produces a single product which is sold at $10 a unit. Suppose further that all variable costs associated with the production of one unit of the product amount to $5 a unit. The latter cost is assumed to be constant for all levels of production. Assume that besides these variable costs the firm has to pay fixed costs of $5,000 a month irrespective of the level of production.[1]

[1] The Encyclopedia Brittanica credits Dionysus Lardner (1850) with being the first to make distinctions between fixed and variable costs and it credits the idea of profit charting and flexible budgeting to John Manger Fells (1903). W. Rautenstrach and E. Knceppel are also mentioned as early advocates of profit-volume planning and breakeven charting. See also Solomons,1952, p.34 ff. and the reference to C.A. Guilbault on pp. 59 ff in Garner, 1954.

On these assumptions we can prepare the following profit graph and find that the firm will break even (i.e. achieve neither profit nor loss) if they can sell 1,000 units a month of the products,

or in other words, if they can achieve a total sales of \$10,000.

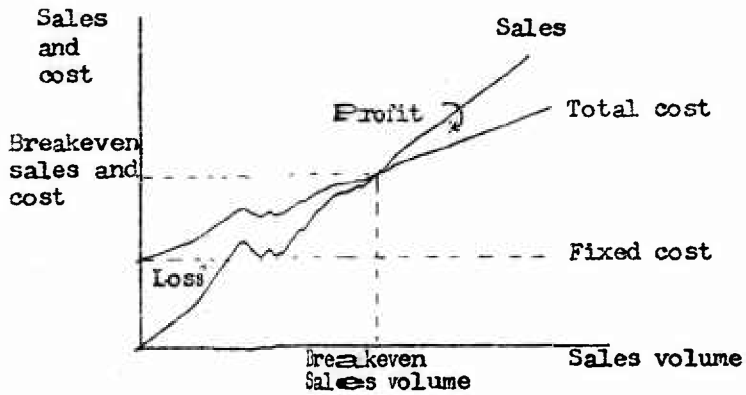


Figure II-1

A Breakeven Chart

This graphical solution can also be verified algebraically. If we let x stand for the breakeven sales volume in physical units, total revenue is given by $10x$ whereas total costs are given by $5x + 5{,}000$. In order to obtain a breakeven point, we set total revenue equal to total costs, i.e.

(2.1) $$10x = 5x + 5{,}000$$

or $$5x = 5{,}000$$

$$x = 1{,}000$$

and this value for x (sales volume) is the unique breakeven point.[1]

---

[1]Refer to Kohler [1952] for a definition and a brief introduction to breakeven chart and breakeven point. More detailed discussions may be found in the literature cited in the bibliography. Note that in the ordinary breakeven analysis the sales volume and the production volume are assumed to be equal, thus avoiding issues on dynamic production planning. We shall develop our analysis based on the same assumption, but extensions to multiple period breakeven analysis can readily be effected by such techniques are ones discussed in Section 6 of Chapter VI.

The analysis is quite straightforward. However, let us pause for a moment and ask ourselves (1) why do we worry about a breakeven point and (2) what is gained by knowing that a breakeven sales volume is 1,000 units?

The answer to the first question may seem to be very simple, i.e. people who operate a business firm (managers) are deeply concerned with whether or not the business has reached a profitable region or not. For managers, there is a qualitative difference between profit and loss, although arithmetically the two are on the same continuous scale. For them, the difference between $1,000 profit and $1,000 loss seems to be much greater than the difference between $10,000 profit and $8,000 profit or the difference between $10,000 loss and $8,000 loss.

The above characterization, if accepted, may also be extended in a variety of ways. Notice that this no-profit no-loss point is not the only one which can have such a psychological "qualitative difference" property. A manager of a firm that has been operating profitably may be concerned with achieving at least the profit level of the preceding year rather than simply achieving a profit. If the profit of the preceding year was, say, $10,000, then a manager might be much more concerned with the difference between $11,000 profit and $9,000 profit than the difference between $5,000 profit and $3,000 profit. In this situation, the $10,000-profit point is as significant as, or more significant than, the zero-profit point.

Such psychological-institutional points which are often of significance in business operations may be referred to as a manager's goals, in order to distinguish them from other commonly assumed objectives like maximum profit, maximum growth, etc. Such goals may be imposed externally (e.g. by boards of directors) or they may be imposed internally, as when a manager's own psychology leads him to impose such goals which are both a constraint and an objective.[1] Once a goal is set, it is assumed that some tension

[1] See, e.g. Charnes and Cooper, 1961, Appendix B.

will be created for a manager that causes him to become concerned about achieving the goal. Conversely when a goal is achieved, the tension is reduced and the desire for further actions is also reduced unless, perhaps, further stimulated by another higher goal.

This kind of goal striving can be made to accomodate other, narrower, concepts like the profit maximization, etc., that are usually assumed to apply in economics and elsewhere. For the moment, however, we may draw certain contrasts. In traditional theory, a manager is supposed to seek a maximum profit. In other words he should be an "optimizer." This should not be equated to the idea of rationality, however, since the requisite degrees of knowledge may be absent as is the case, for instance, when uncertainty is present. It may then be the case that it is rational for a manager to be a "satisficer" and thereby proceed "rationally" towards goals that

he establishes for himself or others, rather than to seek only an optimum point.[1] Such goal-oriented practices may also play

[1]March and Simon (1958, p.140 ff) discuss these satisficing notions in order to contrast them with optimizing properties in terms of a discussion toward "satisfactory vs. optimal standards." Charnes and Cooper (1960), on the other hand, are able to obtain analytical characterizations that relate certain important aspects of both approaches by means of extremal characterizations under conditions of probabilistic constraints.

a crucial role as a managerial device for organizing activities by others via the establishment of tensions or targets needed to "get things done."[1]

[1]Actually, the relationship between a goal and tension is not as simple as this. Uncertainty associated with goal achievement is a major factor influencing the level of tension. However, at this point, it is sufficient to note that such goals give rise to a cyclical fluctuation of tension and qualitatively differentiate points on a continuous scale. Refer to Krech and Gruthfield, 1948, Murphy, 1954, etc., for more detail discussions of the relationship among needs, goals, tension, and motivation.

As already noted and as will be elaborated later in more detail, it is possible to consider optimization, too, as a special kind of goal orientation. Furthermore, even when a manager wants to optimize, he may have several target points which he wants to attain successively during the course of optimization, and, therefore, it may be of interest for him also to know when he has achieved these targets. A goal orientation or analysis may then be valuable even when an optimization is at issue.

By means of these considerations, we can generalize from breakeven to goal analyses without even compromising ourselves or losing anything with respect to optimization. Thus, in this sense the term "goal analysis" or "target analysis" will be used in lieu of the narrower ideas of "breakeven analysis," "optimization," etc. Our concept of goals can, thus, be specialized to breakeven points or optimum points when the latter are of interest.

Some further preliminaries are also in order here if we are to distinguish between breakeven and a more general goal analysis. Once a goal is set, the next step for the manager is to search for means that will help to achieve it.[1] Thus, a manager may

[1]This does not imply that he will forego the possibility of searching for other goals, too, when he confronts causes that make him dissatisfied with the goals he is currently entertaining.

want to break an overall goal down into a set of more operational subgoals that are more or less amenable to organizational manipulation -- e.g., because these subgoals can be related to factors that are under the control of his subordinates.

There are issues of information economy and control, too, which require consideration in this connection. Thus, it may also be of interest for him to have some indicators to tell him when he has reached the goal. For example, it may be easier, quicker, less expensive, etc., for him, or his subordinates, to watch sales volumes than to focus on a net profit figure, which

can be calculated only after taking into account cost factors, variance allowances, and so on.

A breakeven analysis and its more general form, a goal analysis, can make a real contribution to management by translating an overall goal into more <u>operational</u> and <u>measurable</u> subgoals. A simple illustration may be synthesized as follows. Let b be any overall goal, such as the level of sales needed to attain a breakeven point. Let $x_1, x_2, \ldots, x_n$ be the sales levels attainable from each of n products. If we define a functional relation

$$f(x_1, x_2, \ldots, x_n) = b \tag{2.2}$$

which corresponds to all possible mixes which yield the wanted breakeven value, then we can regard each of the variables $x_1, x_2, \ldots, x_n$ as corresponding to subgoals that can be used to secure the overall result (or goal) $f(x) = b$, where $x = (x_1, x_2, \ldots, x_n)$.

This relationship between the goal and a set of subgoals formulated in (2.2) will be analyzed and extended in more detail in this chapter and the next one. Then in Chapter IV we shall attempt to see how such subgoals might be reconsolidated. A major purpose of this analysis is to establish ways for synthesizing indicators for convenient managerial uses in assessing how far or close they are to attaining a specified overall goal level.

2. Single Goal with Single Subgoal

1. Introduction

We proceed now to examining possible ways of synthesizing functions for ascertaining suitable ways of measuring the degrees of goal attainment. Let us start with the case where a goal needs to be translated in terms of only one subgoal. The above graphical analysis associated with (2.1) is taken as a typical example involving one goal (profit level) and one subgoal (sales volume). In this case, we have one goal, b, and one subgoal variable, x, which are related to each other by the following function:

$$f(x) = (10 - 5)x - 5{,}000 = b. \tag{2.3}$$

Here we have a very simple linear relationship between the goal and the subgoal, and we can now submit this relation to further analysis very readily. Let p be unit sales price and c the variable costs per unit. Then if h is fixed cost per period and x represents possible sales volume, we can relate all this to total profit, b, by

$$(p - c)x - h = b \tag{2.4}$$

or, when $p \neq c$

$$x = \frac{h + b}{p - c}. \tag{2.5}$$

At a breakeven point we have, by definition, $b = 0$, so that this further simplifies to

$$x = \frac{h}{p - c} \tag{2.6}$$

If h, fixed cost, is also equal to zero, the breakeven sales volume is always zero. Of course, if $p = c$, this expression (2.5) is not

defined, except we know that if $h > 0$ and $p = c$ then the firm can never break even. If $h > 0$, p has to be greater than c in order to have a non-negative breakeven sales volume. For formal completeness we may also say that if $h < 0$, c has to be greater than p in order to have a non-negative breakeven sales volume, but this case is usually of no interest unless, perhaps, h is interpreted as an initial endowment or subsidy. (See Figure II-2 below.)

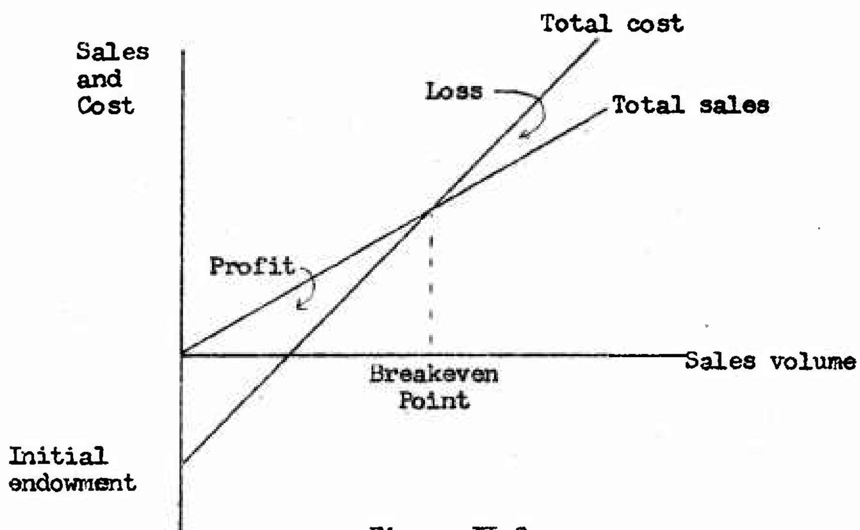


Figure II-2

A Breakeven Chart
(with a negative fixed cost)

2. **Extension to a Piecewise Linear Breakeven Model**

In the above basic breakeven model, it was assumed that the price of the product is constant over the entire range of a feasible production volume, and that the costs of the product consist of only fixed costs and variable costs. We shall extend this basic model to include cases in which the price of the product is no longer constant and its cost schedule involves so-called semi-variable and semi-fixed costs.

First of all, consider the following piecewise linear total revenue curve.

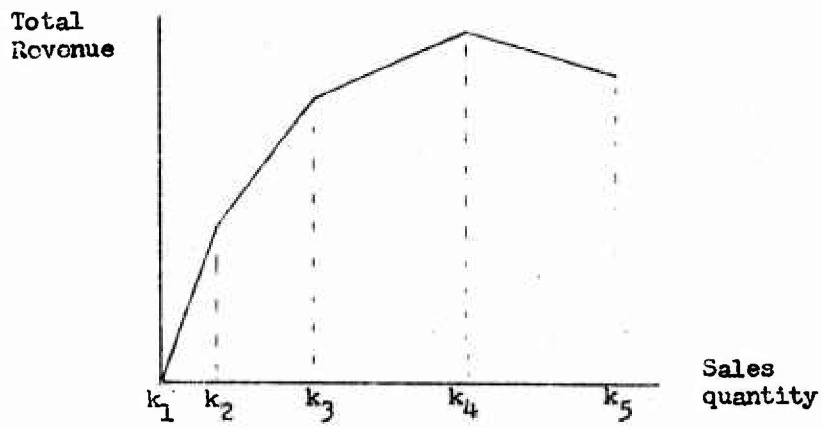


Figure II-3

A Total Revenue Curve

In order to derive analytic expressions for this geometric picture of a total revenue curve, we now proceed as follows. Let r_j be the marginal revenue between the sales volumes k_j and

k_{j+1} $(j = 1, 2, \ldots, m)$ with $k_1 \equiv 0$. Then, in any interval $k_j \leq x \leq k_{j+1}$, we can obtain the total revenue $R(x)$ by means of the relation,

(2.7) $$R(x) = r_1 k_2 + r_2(k_3 - k_2) + \ldots + r_{j-1}(k_j - k_{j-1}) + r_j(x - k_j).$$

This can be rewritten as follows.

(2.8) $$\begin{aligned} R(x) &= r_j x - \sum_{i=1}^{j}(r_i - r_{i-1})k_i \\ &= \sum_{i=1}^{j}(r_i - r_{i-1})x - \sum_{i=1}^{j}(r_i - r_{i-1})k_i \\ &= \sum_{i=1}^{j}(r_i - r_{i-1})(x - k_i), \end{aligned}$$

where $r_0 \equiv 0$ and $k_1 \equiv 0$. This means that we can cover all relevant x values by means of these expressions and the corresponding constraint relations.

We now want to replace this collection of functions and constraints by a single model and so extend the analysis further in this direction. We shall introduce new variables, which are defined and constrained to produce a wholly linear problem.[1]

---

[1]Refer to Charnes and Cooper, 1961, pp. 351 - 55.

This is done as follows. We insert for each range of x a pair of variables x_i^+, x_i^- which are defined and constrained by

(2.9) $$x_i^+ - x_i^- = x - k_i$$

where

(2.10) $$x_i^+ \cdot x_i^- = 0$$

(2.11) $$x_i^+, x_i^- \geq 0 \quad (i = 1, 2, \ldots, m)$$

Observe now that if the x_i^+ and x_i^- $(i = 1, 2, \ldots, m)$ in (2.9) are always chosen in accordance with (2.10) and (2.11) then they will have the characteristic that if $x \geq k_i$ then $x_i^+ = x - k_i$ and $x_i^- = 0$ while

if $x \leq k_i$ then $x_i^- = k_i - x$ and $x_i^+ = 0$. Also these variables x_i^+ have the further desirable property that when $k_j \leq x \leq k_{j+1}$, for any j, then

$$x_i^+ = x - k_i \quad \text{for } i = 1, 2, \ldots, j$$
$$x_i^+ = 0 \quad \text{for } i = j+1, j+2, \ldots, m$$

since $k_0 < k_1 < \ldots < k_j \leq x \leq k_{j+1} < k_{j+2} < \ldots < k_m$. That is, for all preceding $i \leq j$ we will have $x_j^- = 0$ and for all succeeding $i > j$ we will have $x_i^+ = 0$. This permits a further simplification since by using the variable x_i^+ ($i = 1, 2, \ldots, m$) the expression (2.8) which was applicable only for $k_j \leq x \leq k_{j+1}$ can be rewritten in the following form, which is applicable for any x in the range of $0 \leq x \leq k_{m+1}$

(2.12)
$$R(x) = \sum_{i=1}^{m} (r_i - r_{i-1}) x_i^+$$
$$x_i^+ = x - k_i + x_i^-$$
$$x_i^+ \cdot x_i^- = 0$$
$$x_i^+ , x_i^- \geq 0$$

Having thus obtained a general representation for piecewise linear revenue functions, we next turn our attention to the cost side of the problem in order to derive a similar expression for the cost function. Here we should perhaps first emphasize accounting approaches to this topic. In accounting literatures dealing with breakeven analysis, the cost function ordinarily consists of fixed costs and variable costs, -- the first category representing any costs that are constant regardless of the production volume and

the latter varying with volume. Generally accounting analysis proceeds by reference to certain "linearizations." First the variable costs are assumed to vary proportionately with production volume, while fixed costs (also linear) are assumed to be constant, independent of volume. Of course, this is not always adequate, in which event two new categories of cost variations may be introduced under the names "semi-variable" and "semi-fixed."[1]

[1]"Semi-variable" costs are defined as "costs which vary with the production volume but not proportionately." (Kohler, 1952: Anthony, 1960.) However, "semi-variable" is also used to denote costs having fixed and variable components, without implying any non-linear characteristics. (Hill and Gordon, 1959, Lang 1951.) On the other hand, Schlatter defines "semi-fixed" costs as costs that increase abruptly at some critical points and remain constant at other points (Schlatter, 1939).

For our purposes we may consider the latter category to be reflected in terms of certain "jumps" or "discontinuities" of total cost which occur at certain $x = k_j$ values and then relegate the former to certain "kinks" or discontinuities in the derivatives at various k_j values. With this in mind we can separate these two additional categories for specific examination and thereby relate them to our former categorization.

First of all, let us consider the semi-fixed costs, which we represent as a step function in Figure II-4.

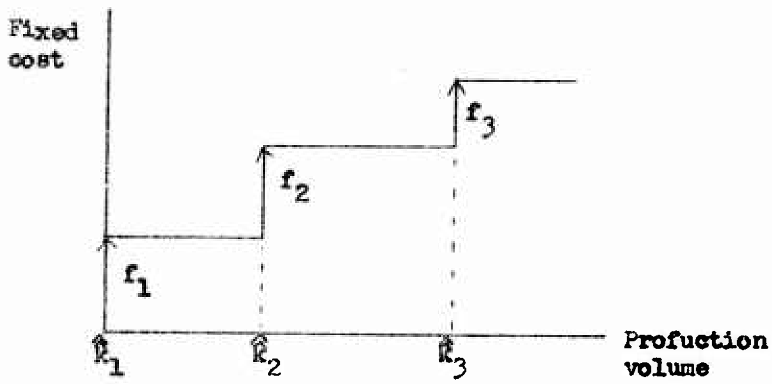


Figure II-4

Semi-Fixed Costs

That is, in contrast to "fixed costs," which are represented by a horizontal line over all x, these "semi-fixed" costs are associated with a series of horizontal line segments which vary only at particular points $x = \hat{k}_j$ and otherwise remain constant.

Next we turn to the category called semi-variable costs and represent them as in Figure II-5. Here again the ordinary variable costs represented by one line which is proportional to x over its entire range is replaced by a set of line segments which have different slopes (proportionality constants) over different ranges $\hat{k}_j \leq x \leq \hat{k}_{j+1}$, $\hat{k}_{j+1} \leq x \leq \hat{k}_{j+2}$, etc.

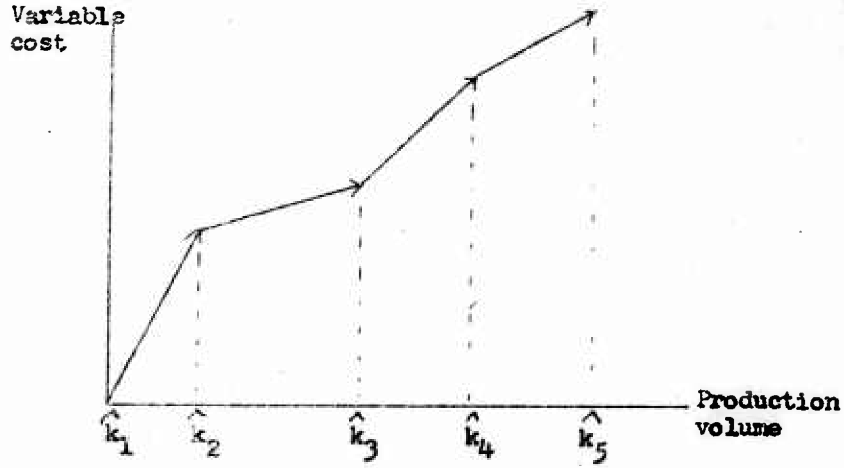


Figure II-5

Semi-Variable Costs

Now we wish to derive a convenient analytic representation for all these costs -- variable, fixed, semi-variable, and semi-fixed. For this we first interpret all costs which are ordinarily considered to be variable in such a way that they are only special instance of a more general category of "semi-variable" costs. Thus, as in the above diagram, any curve of total variable costs can be described completely by a set of three factors which we associate with each line segment as a component of total variable cost. These are (1) the production volume at which the line segment (a variable cost component) starts ($\hat{k}_j$), (2) the slope of the segment (v_i), and (3) either (i) the length of the range in which the proportionality holds (or the "proportionate range"), $d_i = \hat{k}_{i+1} - \hat{k}_i$ or (ii), alternatively,

the production volume at which the segment ends ($\hat{k}_{i+1}$). Thus if there is no semi-variable behavior present, a variable cost is represented by a cost function which has only one slope, v, which holds over the entire range of x. In that case we have our total variable cost completely described by specifying v_1 and d_1.

Now we continue our synthesis of a total cost function for one-product (or homogeneous product) analysis by reference to a consideration of fixed and semi-fixed component of cost. As we already saw, we may define a fixed cost component as the cost increment that enters at a given production volume ($\hat{\hat{k}}_i$) in a given amount (f_i) and then remains constant over a finite range of $x \geq \hat{\hat{k}}_i$. This is in contrast with a variable cost component that also commences at a given production volume ($\hat{k}_j$) but then continues to add a given amount of increment (v_j) per unit increment in the production volume over its "proportionate range" (d_j). Combining these two ideas we can then extend the ordinary accounting definitions of "fixed" cost and "variable" cost to include these "semi-fixed" and "semi-variable" considerations. Hence, by replacing these "fixed vs. variable" classifications over <u>all</u> ranges with the variations that appear over specific ranges we are in a position to identify fixed, semi-fixed, variable, and semi-variable components of total cost according to whether jumps or kinks appear in particular intervals.

Of course, there may be approximation issues, too, as we shall see later, and one that can be usefully singled out occurs when our procedures of model construction or computation cause us to use various artifacts in order to achieve practical implementation of our results. Consider, for instance, a fixed cost component which occurs at $\hat{k}_i$ with the amount f_i and a variable cost component which commences at the same production volume $\hat{k}_i$ with increment v_i per unit of the production volume for the proportionate range d_i. The range, d_i, is assumed equal to f_i/v_i, so that the total cost increment from both cost components are the same. This can be given diagrammatic clarification as in Figure II-6 where a fixed cost and a variable cost component are represented by the two vectors, (F) and (V), respectively.

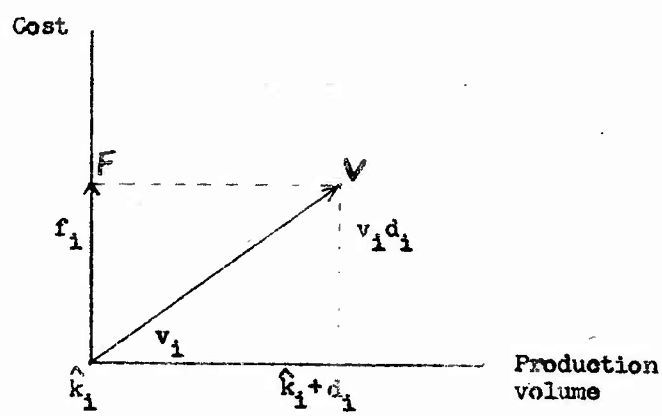


Figure II-6

A Fixed Cost Component and
A Variable Cost Component

Note that in general we need three factors in order to specify a variable cost component (i.e., $\hat{k}_i$, v_i, and d_i) whereas we need only two factors in order to specify a fixed cost component (i.e., $\hat{k}_i$ and f_i.) As this representation makes clear we can regard a fixed cost component as a special case of variable cost components in which d_i is set equal to zero and v_i is replaced by f_i. This is only a formalism, however, since we need to note that there is also a qualitative mathematical difference between v_i and f_i even though both can be interpreted as marginal (incremental) costs as we shall see in the next subsection. In particular v_i has the ordinary concept of calculus derivative as its precise correspond whereas f_i does not. Alternatively, if we proceed by means of the definition of marginal costs relative to the slope of the total cost function -- assumed to be everywhere defined -- then at the fixed-cost jumps we will have an infinite slope.

To accomodate this way of representing a total cost function we can introduce a suitable device as follows.[1] For instance,

[1]Alternatively, we may use so-called diophantine (integer) linear programming techniques. Cf., e.g., Charnes and Cooper, 1961, Ch. XVIII.

we may introduce a vector that commences at $\hat{k}_i$ with unit increment $v_i = f_i/\varepsilon$ over the proportionate range of ε, where ε is a small positive number. The total contribution of this variable cost way of representing the semi-fixed cost jump is then f_i since $v_i = (f_i/\varepsilon)\,\varepsilon = f_i$. (Refer to Figure II-7 below.)

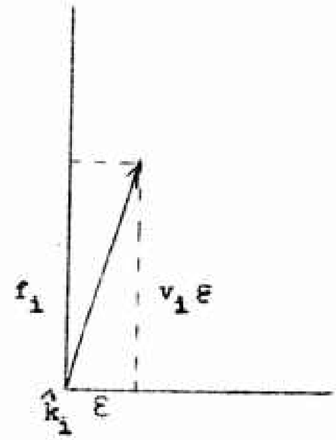


Figure II-7

A Variable-Cost Way of Representing A Fixed Cost Component

By setting $\varepsilon > 0$ sufficiently small, we can, in this way, approximate as close as we please the behavior of total costs by only variable cost components.

There are certain issues of optimality that need to be addressed but for purposes of our breakeven approach we do not need to examine them in detail.[1] For our purposes, then, we may say that any cost

[1] Refer to Charnes and Cooper, 1961, Chapter X and XVIII.

curve can be represented by a suitably arranged set of variable cost components, each one of which is specified by a set of v_i, $\hat{k}_i$ and d_i or alternatively by a set of v_i, $\hat{k}_i$ and $\hat{k}_{i+1}$ $(i = 1, 2, \ldots, n)$ as shown in the following diagram.

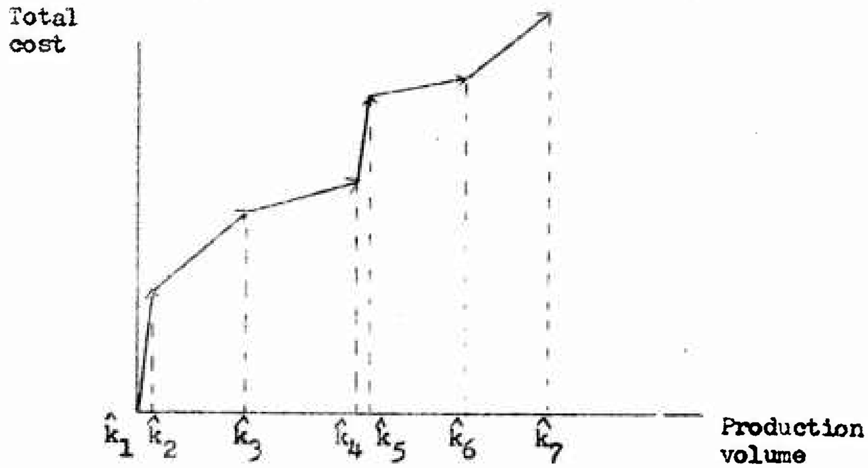


Figure II-8

Approximated Total Cost Curve

Once the cost curve is represented by a piecewise linear function, the total cost $C(x)$ which is a function of the production volume x can be derived in an analogous way as in the total revenue function, i.e. for $\hat{k}_j \leq x \leq \hat{k}_{j+1}$

(2.13) $$C(x) = v_1\hat{k}_2 + v_2(\hat{k}_3 - \hat{k}_2) + \ldots + v_j(x - \hat{k}_j)$$
$$= v_j x - \sum_{i=1}^{j} (v_i - v_{i-1})\hat{k}_j$$
$$= \sum_{i=1}^{j} (v_i - v_{i-1})x - \sum_{i=1}^{j} (v_i - v_{i-1})\hat{k}_i$$
$$= \sum_{i=1}^{j} (v_i - v_{i-1})(x - \hat{k}_i)$$

where $\hat{k}_1 \equiv 0$ and $v_0 \equiv 0$

By the use of the variables x_i^+ and x_i^- described in deriving the total revenue curve, we rewrite (2.13) for a general case in which x is allowed to vary over $0 \leq x \leq \hat{k}_{n+1}$ as follows:

(2.14)
$$C(x) = \sum_{i=1}^{n} (v_i - v_{i-1})\hat{x}_i^+$$
$$\hat{x}_i^+ = x - \hat{k}_i + \hat{x}_i^-$$
$$\hat{x}_i^+ \cdot \hat{x}_i^- = 0 \qquad (i = 1, 2, \ldots, n)$$
$$\hat{x}_i^+, \hat{x}_i^- \geq 0$$
$$0 \leq x \leq k_{n+1}$$

Thus, we obtain approximate expressions for the total revenue curve and the total cost curve, from which the net profit function can be obtained easily by subtracting the latter from the former. The resulting expression can be loaded to a linear programming formulation as follows by introducing two additional variables y^+ and y^- with the condition that $y^+ \cdot y^- = 0$ and $y^+, y^- \geq 0$.

(2.16)

Minimize $\qquad y^+ + y^-$

Subject to

$$\sum_{i=1}^{m} (r_i - r_{i-1})x_i^+ - \sum_{i=1}^{n} (v_i - v_{i-1})\hat{x}_i^+ - y^+ + y^- = 0$$
$$x - x_i^+ + x_i^+ \qquad = k_i \quad (i = 1, 2, \ldots, m)$$
$$x \qquad -\hat{x}_i^+ + \hat{x}_i^- = \hat{k}_j \quad (j = 1, 2, \ldots, m)$$
$$x \leq k$$
$$x, x_i^+, x_i^-, \hat{x}_i^+, \hat{x}_i^-, y^+, y^- \geq 0$$

The solution to (2.16) gives us all the breakeven points. Even though the bilinear constraints $x_i^+ \cdot x_i^- = 0$, $\hat{x}_i^+ \cdot \hat{x}_i^- = 0$, and $y^+ \cdot y^- = 0$ are not shown in the above formulation, the simplex method[1] of solving a linear programming problem guarantees that these (bilinear) constraints

---

[1] I.e., standard versions of this method. In other cases there may be need for recourse to modifications in use of the simplex method, as are available, for instance, by means of the so-called "restricted basis entry procedures" that were developed by A. Charnes and W.W. Cooper. See Ch. X, op. cit.

be satisfied in a solution. We shall discuss this topic in more detail in Section 4 of this chapter.[1]

[1]The reader is again referred to Chapters X and XVIII in Charnes and Cooper, 1961, for a discussion of such issues as local vs. global optima, modifications in the ordinary simplex algorithm routine, etc.

3. Justification for a Piecewise Linear Breakeven Model

Before leaving this section, we would like to raise some arguments on the issues of approximations in order to find some justifications for the use of a piecewise linear breakeven model.

We have used the term an "approximated" cost curve for such a piecewise linear cost curve like one given in Figure II-8. Let us, however, compare such a piecewise linear cost curve with a continuous and everywhere differentiable cost curve of the type normally dealt with in economic theories. Clearly in the real world, we cannot produce 1.5 units of automobiles, 3.7 units of refrigerators, 0.4 units of air-compressors, etc. In some cases, it may be physically possible to produce a fractional unit of product, such as .001 gallon of gasoline, but it would never be of interest of managers in the real world to know the effect on the total cost by such a small increment in the production volume, not to mention any infinitensimal change in the production volume. We may therefore say that the real cost function which we are

interested in is a discrete step function. It is, of course, perfectly legitimate to approximate such a step function by a continuous and differentiable function for the purpose of economic analyses. However, we want to emphasize that such a piecewise linear cost function as given in Figure II-8 is intended to approximate a real cost curve but is not intended to approximate a continuous and differentiable cost curve simply because the latter is normally used in economic analyses. As a matter of fact a piecewise linear cost curve can be a precise reflection of a real cost curve, at least over realistic ranges of assumed production alternatives, whereas a differentiable function need not have this property *ipso facto*. This point seems to be necessary because some of the accounting literature appears to be unduly apologetic here, perhaps in deference to others, such as economists, who are also concerned with the costing problems of business management. But in this connection it may be recalled that until the advent of linear programming, an assumed differentiability was needed to obtain access to the differential calculus as was needed to discuss such issues as optimization, etc.

Professor P. A. Samuelson has stated this all very nicely as follows:

> It is curious to see the logical confusion into which many economists have fallen. The primary end of economic analysis is to explain a position of minimum (or maximum) where it does not pay to make a *finite* movement in any direction. Now in the case that all functions are continuous, it is possible as a means towards this end to state certain *equalities* on differential coefficients which will insure that certain *inequalities*

> will hold for finite movements. It is no exaggeration to say that infinitesimal analysis was developed with just such finite applications in view. Unfortunately, the means have become confused with the ends, and so conventions and artifices are continously sought in order to be able to make statements concerning marginal equivalences...[1]

---

[1] Samuelson, 1947, p. 75.

But with the advent of linear programming and related techniques it is no longer necessary to follow the dictates of the calculus. This can be done, moreover, without undue concern about the imputations which are also often wanted and which the differential calculus supplies when differentiability is assumed. Notice, for instance, as in Figure II-7 any cost component can be made attributable to a unit of product, e.g., k_j^{th} unit, even when an "infinitesimal change" in the production volume causes a discrete finite jump. Hence, as this case shows, we can retrieve our underlying costs as wanted and hence we remain on solid grounds here, too, when we lump the fixed cost component and a variable cost component together in one total cost function via a piecewise linear analysis.

A similar approximation argument also holds for the total revenue curve. Here, too, it is not clear that the total revenue curve must necessarily be a continuous and differentiable function and, furthermore, it is perhaps a more common situation for the relationship between the total revenue and the sale quantity to be linear if small ranges of the sales quantity are only being considered.

Of course, many other ideas in economic analysis have been developed around the assumption of differentiability. The notion of "demand elasticity" is a case in point. But evidently such an idea does not fit readily into the kinds of discontinuities involved in our approximations. Although these ideas may be useful, it does not follow that we should distort our picture of reality because of this, or abandon other aspects of its convenience only for the sake of obtaining a meaningful elasticity. The lack of a meaningful elasticity of demand, as ordinarily defined, can be supplemented, as we shall soon see. But first we can, again, quote Professor Samuelson for immediate perspective.

> There is perhaps some usefulness of the concept of elasticity of demand as giving an indication of the qualitative behavior of total revenue, but even this is only the consequence of neglecting to deal with total revenue directly.[1]

---

[1]Samuelson, 1947, p. 125, Footnote 1.

Note that marginal revenue coefficients, r_i $(i = 1, 2, ..., m)$ in our total revenue curve given in Figure II-4 change their value only at critical points, k_i $(i = 1, 2, ..., m)$, at which elasticity of demand fails to exist. However, we may define a new coefficient which we shall call "marginal revenue head,"[1] denoted by ρ, as meaning

---

[1]The term "head" was borrowed from hydrodynamics.

---

the difference between the two marginal revenue coefficients r_i and r_{i-1} at the critical sales volume k_i and zero everywhere else.

More precisely, we define

(2.17) $$\rho = r_i - r_{i-1} \quad \text{at } k_i \ (i = 1, 2, \ldots, m)$$
$$\rho = 0 \quad \text{at any other point.}$$

Then, by knowing marginal revenue heads at critical points, the total revenue curve can be completely described. If elasticity of demand is intended to indicate some behaviors of the total revenue curve, the same role can be played by marginal revenue head defined above. (Consider positive, zero, or negative marginal revenue heads as a characterization of the total revenue curve.) Of course such a marginal revenue head is not dimensionless, but if dimensionless is what we want,[1] we can use "relative marginal revenue head" defined as marginal revenue head divided by total revenue at that sales quantity.

---

[1]Refer to the following quotation from Samuelson, 1947, p. 126 Footnote 4.

> Actually it is a little misleading to say that an elasticity expression is necessarily "without dimension." For take any absolute derivative, such as dx/dp, which is certainly not dimensionless, involving as it does the dimensions [output times output divided by value.] Even though it has dimensions, it is still the elasticity of some expression. Thus if
>
> $$x = f(p),$$
>
> and
>
> $$y = s\ (q),$$
>
> where
>
> $$y = e^x, \quad q = e^p,$$
>
> then
>
> $$\frac{Ey}{Eq} = \frac{dx}{dp}.$$

The usual role of an ordinary elasticity coefficient in price theory is derived from the fact that it provides a convenient way of summarizing the quantity response to any small variation in price. But this in turn can be related to marginal revenue by

(2.17a) $$\text{Marginal Revenue} = r = \frac{d(px)}{dx} = p + x\frac{dp}{dx} = p(1 + \frac{x}{p}\frac{dp}{dx}) = p(1 - \frac{1}{\eta})$$

where p is the price and η is the elasticity of demand. That is,

(2.17b) $$\eta = \frac{p}{p - r}.$$

In a piecewise linear revenue curve, the total revenue is, of course, represented by a set of line segments as shown in Figure II-3. Thus total revenue may be given representation as a linear function of the sales quantity, x, where x lies between k_j and k_{j+1}, as follows:

(2.17c) $$R(x) = px = r_j x + h_j \qquad \text{for } k_j \leq x \leq k_{j+1}.$$

All symbols have been defined previously except h_j which is the intercept value attained when the j^{th} linear segment is extended to its intersection with the vertical axis.

From (2.17c), we have

(2.17d) $$x = \frac{h_j}{p - r_j}.$$

Within the interval $k_j \leq x \leq k_{j+1}$, the function (2.17d) is differentiable in p and so we can write

(2.17e) $$\frac{dx}{dp} = -\frac{h_j}{(p - r_j)^2} = -\frac{x^2}{h_j}$$

But now, by definition,

(2.17f) $$\eta = -\frac{p}{x}\frac{dx}{dp} = \frac{px}{h_j}$$

for $k_j \leq x \leq k_{j+1}$. In classical economics all functions are generally assumed to be differentiable as required. Hence we can proceed on this assumption and differentiate (2.17f) with respect to x and obtain

(2.17g) $$\frac{d\eta}{dx} = \frac{1}{h_j}\frac{d(px)}{dx} = \frac{r_j}{h_j},$$

with r_j/h_j fixed so that $d\eta/dx$ is constant within any such interval.

Evidently, then, within any interval $k_j \leq x \leq k_{j+1}$, we can integrate (2.17g) over x and obtain η as required. Of course this may be done for any interval and each applicable η can be obtained in this fashion. Also, from (2.8) and (2.17f) we obtain

(2.17h) $$\eta = \frac{1}{h_j}\sum_{i=1}^{j}(r_i - r_{i-1})(x - k_i) \qquad \text{for } k_j \leq x \leq k_{j+1},$$

where $r_0 \equiv 0$ and $k_1 \equiv 0$. We can thus work directly with marginal revenue, r_i, or with marginal revenue head, $\rho_i = r_i - r_{i-1}$, which will usually be the focus of the analysis -- especially since the latter is a dual evaluator under simple revenue maximization[1] --

[1] I.e., dual evaluators associated with the first j constraints in the following linear programming problem:

$$\text{Maximize} \quad R(x) = \sum_{i=1}^{m} (r_i - r_{i-1})x_i^+$$

$$(2.17i) \quad \text{Subject to} \quad x - x_i^+ + x_i^- = k_i \quad (i=1,2,\ldots,m)$$

$$x \leq k_{m+1}$$

$$x, x_i^+, x_i^- \geq 0$$

when an optimal solution to (2.17i), x^*, lies in $k_j \leq x^* \leq k_{j+1}$.

for it will ordinarily be of more immediate managerial and accounting interest without thereby prejudicing our access to other theoretical constructs, like demand elasticities, when the latter have any potential value or use. Although we have here obtained a precise correspondence between our piecewise linear model and an analytical construct from economic theory it is possible that such a precise correspondence may not be forthcoming in all other areas. On the other hand it should be possible to compensate for this in any case of business - managerial interest by first dropping back to basic ideas and then utilizing some ingenuity to attain a practically useful correspond.

This is about as far as we intend to go in this direction. From our standpoint the mathematical methodologies like linear programming make it unnecessary to stay within the strictures prescribed by the classical calculus methods. Switching to the former approaches opens certain prospects for obtaining closer contacts with the usages and procedures of accounting. On the other hand before effecting such a switch it is desirable to check in order to see whether any ideas of value are also lost in the process. Thus in the first part of this chapter we provide an illustrative development that indicated how a linear programming form could be given to some of the current practices in accounting when they are suitably extended. In the immediately preceding discussion we then showed how even such ideas as the elasticities used for demand analysis in economics might also be used at least with respect to important possible applications they might have in management.

3. Single Goal with Multiple Subgoals: Analyses by Generalized Inverses

At the end of Section 1, we defined a goal analysis as a way of deriving a set of subgoals, $x_1, x_2, \ldots, x_n$ that satisfies

(2.2) $$f(x_1, x_2, \ldots, x_n) = b$$

where b represents a given goal. So far we have considered only cases -- linear or non-linear -- that involve one variable as a possible subgoal. In this section, we shall be concerned with cases that involve multiple subgoals while assuming that the function relating the goal and the subgoals is linear. More precisely, we assume that the overall goal can be suitably expressed as a linear combination of n subgoals, $x_1, x_2, \ldots, x_n$. That is, we assume

(2.18) $$f(x_1, x_2, \ldots, x_n) \equiv a_1x_1 + a_2x_2 \ldots + a_nx_n = b$$

where $a_1, a_2, \ldots, a_n$ are any real numbers so that f is a real linear functional.

In order to simplify the notation, we use the ideas of matrix albebra. Let x be a column vector with components $x_1, x_2, \ldots, x_n$ and let a be a row vector with components $a_1, a_2, \ldots, a_n$. Then, the equation (2.18) is represented equivalently by

(2.19) $$ax = b.$$

We shall call any combination of values for $x_1, x_2, \ldots, x_n$ that satisfies (2.18), or equivalently any x that satisfies (2.19), a solution subgoal, or simply a solution. If there is more than

one x which satisfies (2.19) then we shall call all such solutions collectively a set of solution subgoals, or simply a solution set.

For simplicity, we consider the case where a is a two-component row vector and x is a two-component column vector. Suppose a firm sells two kinds of products, Product 1 and Product 2, and that the contribution to overhead and profit (i.e. selling price less variable costs) per unit of Product 1 and Product 2 are \$1 and \$.5, respectively. Suppose further that the goal in this problem is to earn a profit of \$1 over and above fixed costs of \$1.50. I.e., the goal is to earn \$2.50 (= \$1.00 + \$1.50), thereby covering the fixed cost of \$1.00 and contributing \$1.50 to profit.

We next want to interpret the goal in terms of subgoals that are more immediately operational relative to the categories, Product 1 and Product 2. Let x_1 and x_2 be the sales volume of Product 1 and Product 2, respectively, and let it be supposed that these sales volumes are more operational than the profit variable itself.

Any combination of x_1 and x_2 which satisfies the following condition will exactly attain the given goal:

$$x_1 + .5x_2 = 2.5.$$

or, by using vector notation,

(2.20) $$(1, .5)x = 2.5$$

where x is a two-component column vector for x_1 and x_2. Since negative x_1 or x_2 has no meaning in this case, both x_1 and x_2 have to be non-negative. However, in order to analyze the mathematical

implication of the problem, let us ignore this non-negativity constraint until a later stage of the discussion.

We now want to obtain the solution set to (2.20). We shall do so by first introducing the idea of a transformation. By a transformation we mean any operation -- e.g., a mapping -- that relates points (or vectors) in one space (called the domain of the transformation) to points in another space (called the range of the transformation.) If a transformation is linear, then, it is uniquely associated with a matrix and vice versa, -- i.e., to each linear transformation there corresponds one and only one matrix, and vice versa.

In our example, the 1 x 2 matrix (1, .5) uniquely defines a transformation by which a point in a two-dimensional space is transformed into a point in a one-dimensional space. Then, our problem is equivalent to a geometrical problem of finding the set of the points in the two-dimensional space which are transformed into the point, 2.5, in the one-dimensional space under the transformation defind by the matrix (1, .5).

If a matrix A has an ordinary inverse A^{-1}, then we have a unique solution x to $Ax = b$ which is given by $x = A^{-1}b$. Geometrically, this means that the transformation defined by A transforms x into b, whereas the transformation defined by A^{-1} transforms b precisely back into x. Therefore, if the matrix in our problem of goal analysis has an ordinary inverse, then

the solution subgoal which attains the given goal level can be obtained by using the ordinary inverse of the matrix.

However, only special kinds of matrices, i.e., non-singular (square) matrices, have ordinary inverses, and we cannot limit our goal analyses to the cases which involve only non-singular matrices. Therefore, for our goal analyses, a more general idea of inverse of a matrix is needed.

We shall now introduce the idea and the uses of the so-called generalized inverse [1] of a matrix in our goal analyses

---

[1] Developed by Moore, 1920, and independently by Bjerhammer, 1951, and Penrose, 1955.

---

for various reasons. First, every matrix (singular or non-singular, square or rectangular, zero or non-zero matrices) has its own unique generalized inverse. Thus, we need not be concerned with questions of existence and uniqueness. Moreover, when a matrix has the ordinary inverse (i.e., when the matrix is a non-singular square matrix), the generalized inverse of the matrix is identical with its ordinary inverse, and hence we need not to worry about the uniqueness of the inverse in this respect, too. Second, a generalized inverse incorporates the essential idea of a "solution," which is what we want, and dispenses with the property of uniqueness of a solution, which is not always convenient for goal analysis any way. Finally, a generalized inverse has a so-called "least square property" which is useful for goal analysis in which incompatible

multiple goals are involved as we shall see in a later section.

Even though it is not the main purpose of this chapter to analyze the mathematical properties of a generalized inverse, we feel it useful for readers to have some ideas about how these mathematical properties used in goal analysis in this chapter are derived. Therefore, we have attached in Appendix A an analysis of the generalized inverse by using simple examples.

For an n-component non-zero row vector a, the generalized inverse of a, denoted by $a^{\dagger}$, is an n-component column vector given by $(\frac{1}{aa^*})a^*$, where a* is the transpose of a.[1] In our example,

---

[1] Refer to Section 6 of Appendix A for a computation method of a generalized inverse. See also Charnes, Cooper and Thompson, December, 1962.

---

the generalized inverse of (1, .5) is

$$(1, .5)^{\dagger} = \frac{1}{(1, .5)\binom{1}{.5}}\binom{1}{.5} = \frac{1}{1.25}\binom{1}{.5} = \binom{.8}{.4}.$$

By using the generalized inverse of a, one of many possible solutions to ax = b is given by

$$(2.21) \qquad x = a^{\dagger}b.$$

The x of (2.21) may be called a <u>particular</u> <u>solution</u> to ax = b, just like in solutions to a differential equation. It has a special property which will be utilized in a later chapter.

However, this is only one of many possible solutions to ax = b. To obtain a more general expression which will give

us <u>all</u> the solutions to $ax = b$, consider the following expression for x.

(2.22) $$x = a^{+}b + w$$

where w is an n-component column vector. Since $a^{+}b$ is a constant vector, it follows that for any choice of w, x is uniquely determined and *vice versa*, i.e., there is one-to-one correspondence between x and w. Moreover the condition $ax = b$ is satisfied if and only if $aw = 0$, since,

(2.23) $$ax = aa^{+}b + aw = b + aw$$

The set of all vectors w which satisfy $aw = 0$ is called the <u>null space</u> of a, denoted by $N(a)$. Therefore, <u>all</u> the solutions to $ax = b$ can be obtained by adding $a^{+}b$ to each vector in the null space of a.

All the vectors in the null space of a can be represented by linear combinations of vectors in a null space basis of a. Let a^{o} be an n by $(n-1)$ matrix obtained from $n-1$ column vectors in a null space basis. (For any non-zero row vector a, a null space basis of a consists of $n-1$ linearly independent vectors).[1] Then, all the vectors in the null space of a can be obtained from

[1]A method of computing a null space basis is given in Section 4 of Appendix A.

the expression $a^{o}z$ where z is an arbitrary $(n-1)$-component column vector which acts as a weighting factor of linear combinations of the basis vectors.

This is not, of course, the only way to represent the null space of a. The expression $(I - a^{\dagger}a)z$, where I is an n x n identity matrix, gives a vector in the null space of a by projecting perpendicularly (or by a more technical term, orthogonally) an n-component arbitrary vector z into the null space of a.[1]

---

[1] For more detailed discussion, see Section 5 of Appendix A. In our example, (see (2.20))

$$(I - a^{\dagger}a) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} .8 \\ .4 \end{pmatrix}(1, .5) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} .8 & .4 \\ .4 & .2 \end{pmatrix} = \begin{pmatrix} .2 & -.4 \\ -.4 & .8 \end{pmatrix}.$$

Therefore, any vector x in the null space of a, or equivalently any vector which satisfies $ax = 0$, is given by a suitable choice of the values of the two-component column vector z, as follows

$$x = \begin{pmatrix} .2 & -.4 \\ -.4 & .8 \end{pmatrix} z .$$

Note that any vector thus represented has the form $\begin{pmatrix} -.5 \\ 1 \end{pmatrix}\alpha$. Therefore this expression may be replaced by

$$x = \begin{pmatrix} -.5 \\ 1 \end{pmatrix}\alpha$$

where α is a scalar (a number). That is, any vector x which satisfies $(1, .5)x = 0$ is represented by $\begin{pmatrix} -.5 \\ 1 \end{pmatrix}\alpha$ with a suitable choice of a scalar α. The vector $\begin{pmatrix} -.5 \\ 1 \end{pmatrix}$ is a null space basis of (1, .5) obtained by the method in Section 4 of Appendix A.

This expression, $(I - a^{\dagger}a)z$ $(z \in E^n)$,[1] has some mathematically

[1]"E^n" refers to a real n-dimensional vector space and "ε" means "member of." Therefore, "($z \in E^n$)" means that z is an n-component (column) vector with all of its components real numbers.

interesting properties which cannot be observed in an expression $a^o z$ as discussed in Appendix A. For example, the expression $(I - a^+ a)$ is <u>unique</u> for any a whereas the expression using a null space basis, i.e., $a^o z$ $(z \in E^{n-1})$, is not. We shall, however, use either of the two expressions as convenience for our analyses suggests, noting that any expression for the null space of a may be used in place of $a^o z$ or $(I - a^+ a)z$.

Using the generalized inverse of a and a null space basis of a, the set of all solutions to $ax = b$ is given by

$$x = a^{\dagger}b + a^{o}z \quad (z \in E^{n-1}). \tag{2.24}$$

For example, since the generalized inverse of $(1, .5)$ is $\binom{.8}{.4}$ and a null space basis of $(1, .5)$ may be given by $\binom{-.5}{1}$,[1] the set

[1]See previous footnotes in which the vectors $\binom{.8}{.4}$ and $\binom{-.5}{1}$ are derived.

of all solutions to our example can be given by

$$x = \binom{.8}{.4}b + \binom{-.5}{1}z \tag{2.25}$$

where b is a desired profit level (plus fixed costs) and z is an arbitrary scalar. I.e., here $z \in E^1$

In preparing the expression for the set of all solutions to $ax = b$, as above, it should be noted that we do not need to know the value of the goal, b, beforehand. Both $a^{\dagger}$ and a^{o} can be calculated independently of b. Then, once the goal level is specified, we can immediately get the set of all solutions for subgoals that exactly attain the given goal. Thus by the use of generalized inverses, the goal is directly translated into more operational subgoals.

Depending upon the nature of subgoals in goal analysis, we may want to avoid negative values for subgoals. For example, in our problem of a profit goal attainment, negative sales volumes do not have any meaning. In such cases, we attach the non-negativity condition to the expression for x; i.e.

(2.26) $$x = a^{\dagger}b + a^{o}z \geq 0$$

or

(2.27) $$a^{o}z \geq - a^{\dagger}b$$

This means that the non-negativity condition for x is now translated into constraints on otherwise arbitrary variables in the expression for the null space of a. For example, the non-negativity constraint on x in (2.25) is converted into the constraint on z as follow:

(2.28) $$\begin{pmatrix} -.5 \\ 1 \end{pmatrix} z \geq - \begin{pmatrix} .8 \\ .4 \end{pmatrix} b$$

or

(2.29) $$- .4b \leq z \leq 1.6b$$

for any non-negative b. (If b is negative, no solution exists for $ax = b$, $x \geq 0$, where $a \geq 0$.) Thus, (2.25) together with (2.29) give us the set of all combinations of sales volumes of Product 1 and Product 2 which achieve a specified level of contribution to overhead and profit, b, for any non-negative b.

4. Single Goal with Multiple Subgoals: Analyses by goal programming

In this section, we shall use a special type of linear programming, which is called "goal programming,"[1] to attack the

---

[1]The linear programming application in this section and in the next chapter is a revised version of Charnes, Cooper and Ijiri, 1962. For more detailed discussions of "goal programming," see Charnes and Cooper, 1961, p. 215 ff.

---

same problem of goal analysis as in the previous section.

The basic formulation for goal analysis with a single goal and multiple subgoals was given by the expression (2.19) in the previous section. i.e.

(2.19) $$ax = b.$$

This expression can be rewritten in its equivalent form by using the following linear programming formulation.

(2.30)
$$\begin{aligned} &\text{Minimize} && y^+ + y^- \\ &\text{subject to} && ax^+ - ax^- - y^+ + y^- = b \\ &&& x^+,\ x^-,\ y^+,\ y^- \geq 0 \end{aligned}$$

where y^+ and y^- are non-negative real variables, and x^+ and x^- are n-component non-negative column vectors.[1] In order to satisfy

the non-negativity requirement of variables in a linear programming problem, we split x into x^+ and x^- whose components are all non-negative with the property that

$$x = x^+ - x^- \tag{2.31}$$

Although such a decomposition of x is not unique, we can evidently always write any x as a difference of two non-negative vectors, and, furthermore, we can obtain uniqueness, when desired, by specifying an algorithm with relevant properties -- such as one which assures $x^+ \cdot x^- = 0$[1] -- or else obtain it in other ways.

---

[1]The simplex method of linear programming has this property. This is referred to by Charnes and Cooper (1961) as an algorithmic completion of a model, and the reader may refer to this for more detailed discussion of this topic and related artifacts.

The variables y^+ and y^-, which are called slack variables, are included in the constraint to make the goal analysis more flexible, as we shall see later.

For simplicity of the discussion we shall assume that the variable x is constrained to be non-negative, noting that whenever the nature of the subgoals allows us to have negative components in x we can always express x as a difference of two non-negative variables x^+ and x^- thus causing no trouble in applying the subsequent discussion to such a case. Therefore we shall deal with the following modified version of (2.19) and (2.30), in which x is constrained to be non-negative.

$$(2.32) \qquad \begin{aligned} ax &= b \\ x &\geq 0 \end{aligned}$$

$$(2.33) \qquad \begin{aligned} &\text{Minimize} && y^+ + y^- \\ &\text{subject to} && ax - y^+ + y^- = b \\ & && x, y^+, y^- \geq 0 \end{aligned}$$

The two formulations are equivalent in the sense that any x which

satisfies (2.32) also satisfies (2.33) and vice versa, provided that a solution exists for (2.32). This will be clear from the fact that the minimization process always drive the values of y^+ and y^- down to zero, which is the minimum value that y^+ and y^- can ever take, provided that a solution exists for (2.32).

A solution to (2.32) does not exists if (1) $a \geq 0$ and $b < 0$ or (2) $a \leq 0$ and $b > 0$ since no non-negative x can than satisfy $ax = b$. Of course when a is a vector and b a scalar with no constraint on x, there is no great trouble in specifying things like existence theorems, etc., with respect to solvability. Nevertheless this situation provides us with an opportunity to examine this issue in the context of goal attainment when, in fact, (2.32) has no solution in order to see how it might be reflected in (2.33). Of course $ax = b$ always has a solution unless $a = 0$ and $b \neq 0$. But the condition $x \geq 0$ could not be satisfied if $a \geq 0$ and $b < 0$ or if $a \leq 0$ and $b > 0$. However, (2.33) will always be solvable in non-negative x since we can always choose $x = 0$ and $y^- = b$, $y^+ = 0$ when $b > 0$ or $y^- = 0$, $y^+ = -b$ when $b < 0$. In any case a minimizing solution, when attained, produces a $y^+ + y^-$ discrepancy -- i.e., a non-negative value for y^+ and y^- with $y^+ + y^- > 0$ for the case of no solution -- which is as small as possible so that, conversely, the corresponding x values give a resulting ax which is "as close as possible" to the goal level b.

There is some ambiguity in that the positive or negative deviation between ax and b corresponds to a single variable y which is unrestricted

in sign. Of course the value for any such variable may equally well be represented as a difference of non-negative variables y^+ and y^- so that $y = y^+ - y^-$ with $y^+, y^- \geq 0$ for any value of y whatsoever.

On the other hand, any particular value of y may be represented in many ways by $y^+ - y^-$. For instance, if $y = 2$ then $y = y^+ - y^- = 5 - 3 = 4 - 2$ and so on while if $y = -2$ then $-2 = 3 - 5 = 2 - 4$, etc. We would like, however, to have this ambiguity of choice resolved and this is done by the minimization which insures $y^+ \cdot y^- = 0$ so that at least one of these variables will be zero at a minimum. When this is done then the related uniqueness property resolves the ambiguity. Thus, in the above cases we will have $y = 2 = y^+$ with $y^- = 0$ and $y = -2 = -y^-$ with $y^+ = 0$

To see this point we may consider any x which produces a corresponding discrepancy in amount k. I.e. for any x we have

$$ax - b = k = y^+ - y^-. \tag{2.34}$$

Suppose that we have then $y^+, y^- > 0$. This could not be minimal for the corresponding functional values, however, for writing

$$y^+ = k + y^- \tag{2.35}$$

we see that we could reduce both y^+ and y^- and thereby reduce the value of $y^+ + y^-$. Moreover this reduction could be continued until either or both $y^+ = 0$ or $y^- = 0$ is achieved. Since this is true for any x it must be true for a minimizing x and so we can conclude that $y^+ \cdot y^- = 0$ is a necessary condition of minimization. Thus, at most one of $y^+, y^- \geq 0$ can be positive at a minimum[1].

[1]For an alternate proof see Charnes and Cooper, 1961.

The linear programming problem (2.33) is an example of so-called "goal programming." This is distinguished from other types of linear programming problems by virtue of the fact that at least one of the constraints is incorporated in the objective function in such a manner that it becomes a part of the objective for maximization or minimization. By using this goal programming formulation, our example of profit goal attainment problem is now formulated as follows:

(2.36) Minimize $y^+ + y^-$

Subject to

$$x_1 + .5x_2 - y^+ + y^- = 2.5$$

$$x_1,\ x_2,\ y^+,\ y^- \geq 0$$

We obtain two basic solutions to (2.36), i.e. $x_1 = 2.5$, $x_2 = 0$ and $x_1 = 0$, $x_2 = 5$, from which we obtain the set of all solutions by convex combinations of the two basic solutions. That is, for each $0 \leq \theta \leq 1$ we obtain a solution x by means of the expression

(2.37) $$x = \theta \binom{2.5}{0} + (1 - \theta)\binom{0}{5} = \binom{2.5\theta}{5 - 5\theta}$$

which are called "convex combination" of these two basic solutions. Note that the first basic solution ($x_1 = 2.5$, $x_2 = 0$) corresponds to the solution obtained from (2.25) by setting b = 2.5 and z = -1 (i.e., the lower bound for z given in (2.29)), and the second basic solution ($x_1 = 0$, $x_2 = 5$) corresponds to the solution obtained from (2.25) be setting b = 2.5 and z = 4 (i.e., the upper bound

for z given in (2.29)). It will also be noted that any convex combination of the two basic solutions can be obtained by a suitable choice of z in the range specified by (2.29) in which b is set equal to 2.5. Thus the set of x's determined by (2.36) is identical with the set of x's determined by (2.25) together with (2.29).

In the next chapter, we shall extend our goal analysis to cases in which a choice of subgoal values has to be made in such a way that it satisfies several constraints on subgoals which arise from the environment of business operations. Then, we shall extend our goal analysis further to cases which involve multiple goals.

CHAPTER III

CONSTRAINTS ON SUBGOALS AND MULTIPLE GOALS

1. Introduction

The purpose of this chapter is to extend our models of goal analysis in order to apply them to more complicated situations in business operations.

In Section 2 of this chapter, we shall examine implications of imposing upon subgoal variables a set of constraints which may come from an environment of the business operations or which may be imposed by a higher echelon in the organization. We shall also analyze cases that involve an "unattainable" goal due to the constraints imposed upon subgoals.

In Sections 3 and 4, we shall examine cases in which multiple goals are involved by means of linear programming and generalized inverses, respectively. We shall especially focus attention on cases in which multiple goals are incompatible. Such an incompatibility may come from the nature of the goals themselves or it may come from the constraints imposed upon the subgoals by which the goals are to be attained. In such cases, ordering and weighting of the goals become important ways of solving the incompatibility of the goals, as we shall see in Sections 3 and 4. After introducing the concept of ordering, preemptive priority factors, etc., we shall note that constraints (imposed upon subgoals) may be considered as a

special case of goals in that, notably, a constraint always assumes priority over any functional used in association with an objective even when the functional contains preemptive priority terms for its coefficients.

2. Constraints on Subgoals

In (2.33) of the previous chapter, we were concerned only with conditions like the following for x, the subgoal variable: (i) to attain the goal -- i.e., choose x to satisfy $ax = b$ exactly -- or (ii) if this cannot be done then come as close as possible to the goal.[1/] Only the constraint $x \geq 0$ was otherwise

1/ Both (i) and (ii) are automatically provided for in (2.33), a topic we shall discuss in more detail later in this chapter.

imposed on the subgoal variable. But, of course, we will often want to impose further constraints on the subgoal variables, x, which may come from the environment of the business operations.

Let us represent a new set of constraints which we now want to impose upon the subgoal variables as follows:

(3.1) $$Bx \leq h,$$

where B is an m x n matrix and h is an m-component column vector. Then to stay in touch with (2.33) we now model a situation in which either the goal is attained in the light of the subgoal constraints or else the minimization process produces an x vector which satisfies all subgoal constraints in such a way as to come as close as possible to the goal level b. The model is as follows:

$$\text{(3.2)}\quad \begin{array}{ll} \text{Maximize} & y^+ + y^- \\ \text{subject to} & ax - y^+ + y^- = b \\ & Bx \leq h \\ & x,\ y^+,\ y^- \geq 0 \end{array}$$

Example 1: Consider the following example taken from the previous chapter. A firm produces two products, Product 1 and Product 2, whose contribution to overhead and profit are \$1 and \$.5, respectively. Suppose that it is the manager's goal to attain the profit level of \$1 over and above the fixed costs of \$1.5, namely to obtain \$2.5 for contribution to overhead and profit. We now want to interpret this goal in terms of more operational subgoals, i.e., sales volumes of Product 1 and Product 2.

In addition, suppose we have only 12 hours and 10 hours of machine time available for Machine 1 and Machine2, respectively, whereas each unit of Product 1 requires 3 hours of Machine 1 and 2 hours of Machine 2 and each unit of Product 2 requires 5 hours of Machine 1 only. For simplicity, let us assume that the ending inventory levels for Product 1 and Product 2 have to be maintained exactly at the beginning inventory levels of the period, thus making the sale volume and the production volume equal. Let x_1 and x_2 be the physical sales volumes (and consequently the production volume) of Product 1 and Product 2, respectively. Then, we can express the problem of this goal analysis by the following linear programming problem.

$$\text{(3.3)}\quad \begin{array}{ll} \text{Minimize} & y^+ + y^- \\ \text{subject to} & \\ & x_1 + .5x_2 - y^+ + y^- = 2.5 \\ & 3x_1 + 2x_2 \leq 12 \\ & 5x_1 \leq 10 \\ & x_1,\ x_2,\ y^+,\ y^- \geq 0 \end{array}$$

We have two basic solutions to (3.3), which are (i) $x_1 = 2$, $x_2 = 1$ and (ii) $x_1 = 0$, $x_2 = 5$. In both cases we have y^+ and y^- both equal to zero. This means that the manager's goal

of a profit level of \$1 over and above the fixed costs of \$1.5 is attainable under the constraints imposed upon the subgoals (the sales and production volumes of the two products) and the goal is just attained if the sales product mix is either (i) or (ii) above or any convex combination of the two as in an example given in (2.37) in the previous chapter.

A number of variations of (3.2) can be made for different types of goal analysis by changing the sign of y^+ and y^- in the functional or by dropping one of them. These shall now be briefly examined before we complicate the problem by introducing multiple goals.

Logically we have the following eight variations of the functional, including one in (3.2), which can be combined with the constraints in (3.2). They are listed in Table III-1. We shall examine each one of them separately. For convenience, let us denote by (3.2 - k) the linear programming problem (3.2) whenever its functional is replaced by one indexed by k, k = 1, ..., 8, the number in the following table which identifies the functional.

Table III-1

Variations of Goal Programming Functional

	Functional: Equivalent minimizing and maximizing objectives — Minimize	or Maximize	Choices of x accord with	Resulting values of y^+ and y^-
(1)	$y^+ + y^-$	$-y^+ - y^-$	Minimize $\|ax-b\|$	$y^+=ax-b$, $y^-=0$ if $ax>b$; $y^-=b-ax$, $y^+=0$ if $ax<b$ and $y^+=y^-=0$ if $ax=b$
(2)	y^+	$-y^+$	Minimize ax-b insofar as $ax>b$	As above, $y^+=0$ if possible or min $y^+>0$, if not.
(3)	y^-	$-y^-$	Minimize b-ax insofar as $ax<b$	As above, $y^-=0$ if possible or min $y^->0$, if not.
(4)	$y^+ - y^-$	$-y^+ + y^-$	Minimize ax	$y^+ - y^- = ax - b$
(5)	$-y^+ + y^-$	$y^+ - y^-$	Maximize ax	$-y^+ + y^- = b - ax$
(6)	$-y^+ - y^-$	$y^+ + y^-$	Infinite solution*	$y^+ + y^- \longrightarrow \infty$
(7)	$-y^+$	y^+	Infinite solution	$y^+ \longrightarrow \infty$
(8)	$-y^-$	y^-	Infinite solution	$y^- \longrightarrow \infty$

* Note: Infinite solutions here mean $x \rightarrow \infty$. In (4) and (5) y^+, y^- can go to infinity but the difference between the two is always constant.

(1) The solutions obtained from (3.2 - 1) minimizes $|ax - b|$.
In these solutions, at least one of y^+ or y^- has to be equal to zero as we have discussed in the previous chapter. This means that regardless of whether $ax \geq b$ or $ax \leq b$, the minimization process searches for the x which minimizes y^+ (= $ax - b$ if $ax \geq b$; = 0 if $ax \leq b$) or y^- (= $b - ax$ if $ax \leq b$; = 0 if $ax \geq b$) whichever is larger.[1/]

1/ I.e., term by term, as given by the constraints associated with each y_i^+ and y_i^-, in the general case, it will (a) minimize on the maximum of y_i^+, y_i^- and (b) equate the other one to zero. This kind of generalized "minimax" is therefore also included in the objective of (3.2 - 1).

---

(2) If the functional (2) is used, the set of solutions, or the solution set, consists of all x's which make $ax \leq b$ ($y^+ = 0$) satisfying $Bx \leq h$, $x \geq 0$, provided that such solutions exist. If such solutions do not exist, the solution set consists of all x's which minimize $ax - b$ ($y^+ = ax - b > 0$).

(3) Similarly, if the functional (3) is used, the solution set consists of all x's which make $ax \geq b$ ($y^- = 0$) satisfying $Bx \leq h$, $x \geq 0$, provided that such solutions exist. If such solutions do not exist, the solution set consists of all x's which minimizes $b - ax$ ($y^- = b - ax > 0$)

(4) The effect of the functional (4) is to minimize ax.
To see this, let

(3.4) $$y = y^+ - y^-.$$

Note that y is unrestricted in sign. Then, (3.2 - 4) is equivalently written as

$$\begin{array}{lll} & \text{Minimize} & y \\ (3.5) & \text{subject to} & ax - y = b \\ & & Bx \le h \\ & & x \ge 0, \end{array}$$

or

$$\begin{array}{lll} & \text{Minimize} & ax - b \\ (3.6) & \text{subject to} & Bx \le h \\ & & x \ge 0. \end{array}$$

Since b is a constant the functional in (3.6) is, for optimization, equivalent to "minimize ax."

Note that in the solutions to (3.2 - 4), the variables y^+ and y^- cannot be determined uniquely since any combination of y^+ and y^- which satisfies $y^+ - y^- =$ minimum$(ax - b)$ can be a solution. However, even though the minimization process does not automatically push one of the variables y^+ or y^- down to zero, when the problem is solved by the simplex method either y^+ or y^- is always set equal to zero in a basic solution.[1/] Therefore, in a solution by the

1/ This is because the y^+ vector and the y^- vector in a simplex tableau are dependent upon each other.

simplex method,

$$\begin{array}{ll} & y^+ = ax^* - b \text{ and } y^- = 0 \text{ if } ax^* \ge b; \\ (3.7) & \\ & y^- = b - ax^* \text{ and } y^+ = 0 \text{ if } ax^* \le b, \end{array}$$

where x^* is a solution which minimizes ax^* satisfying $Bx \le h$, $x \ge 0$.

Hence the simplex algorithm resolves the ambiguity footnoted at the

bottom of Table III - 1 so, then, a positive y^+ may be interpreted as meaning that ax cannot be pushed down to the goal level b, or the goal is unattainable, and that the closest we can come to the goal level b is $b + y^+$, and a positive y^+ then measures the discrepancy or distance to the goal b that is achieved by x in ax. On the other hand, a positive y^- may be interpreted as meaning that ax can be pushed down beyond the goal level b and the closest we can come to b via x is $b - y^-$, so that in this case the distance to the goal level is equal to y^-. If $y^+ = y^- = 0$, this means that the goal level b is exactly the minimum attainable level for ax under the given constraints, and this situation would be the same for all of the cases k = 1, ..., 5 which are thus then seen to be equivalent whenever attainable goals only are involved.

(5) The effect of the functional (5) is to minimize -ax or, equivalently, maximize ax. In exactly the same way as in (4), (3.2 - 5) is equivalent to

$$\text{(3.8)} \qquad \begin{aligned} &\text{Maximize} \quad ax - b \\ &\text{subject to} \\ &\qquad Bx \le h \\ &\qquad x \ge 0. \end{aligned}$$

Here again, y^+ and y^- cannot be uniquely determined since any combination of y^+ and y^- which satisfies $y^+ - y^- = \text{minimum } (ax - b)$ can be a solution. However, as in (4), in a solution by the simplex method, the values of y^+ and y^- are always uniquely determined, satisfying (3.7).

Now consider (6), (7), and (8). Since $y^+ = y^- + ax - b$ from (3.2), y^+ may be increased (or $-y^+$ may be decreased) indefinitely without making y^- negative no matter what x is used for $(ax - b)$. Similarly, regardless of the value of $(ax - b)$, y^- may be increased (or $-y^-$ may be decreased) indefinitely without making y^+ negative. Therefore, there is no finite solution for (3.2 - 6), (3.2 - 7), and (3.2 - 8). This means that y^+ and y^- no longer provide any measure of goal attainment for the subgoal variables, x, since no matter what x is used we will always have an infinite solution or discrepancy when $ax = b$ is satisfied exactly.

We now give an example to illustrate these variations as follows:

Example 2: Consider a division of a firm which has n different kinds of activities (e.g., purchases or production of n different kinds of commodities), whose activity levels are represented by elements in a vector x, here interpreted as subgoal variables for this division. The division has to satisfy the requirements imposed upon the subgoals embodied in $Bx \leq h$.

Suppose that the cost of a unit operation of the i^{th} activity is given by the i^{th} element in a row vector a. If ax is the total cost of the program corresponding to activity levels x, we may consider our goal relative to a given budget b.

Then, the solutions to (3.2 - 1) — (3.2 - 5), or "programs", may be interpreted as follows:

(3.2 - 1): Programs which require exactly the given amount of budget b, or, if there is no such program, ones which require an amount closest to b, resulting in a surplus or a shortage in the budget.

(3.2 - 2): Programs which can be implemented within the given budget b, or, if there is no such program, ones which result in a minimum amount of budget shortage, defined as the negative of a budget surplus.

(3.2 - 3): Programs which exhaust the given budget, as in (3.2 - 2) but here if there is no such program, choose ones which result in a minimum amount of budget surplus defined as the negative of a budget surplus.

(3.2 - 4): Programs which require a minimum amount of budget, with either shortages or surplus permitted en route to the minimum.

(3.2 - 5): Programs which require a maximum amount of budget, with either shortages or surplus permitted en route to maximum.

3. Analysis of Multiple Goals: A Goal Programming Approach

1. Introduction

In this section, we extend our analysis of goals to cases which involve multiple goals.

In the traditional theory of economics, profit maximization has been viewed as the sole objective of a business firm. For the purpose of market analyses in economics this may be a satisfactory assumption,[1] but for our purposes of theorizing control accounting

[1] See, however, Cyert and March, 1963.

it is desirable to have a "representative firm" which is different from the ones used in economic analyses, in order to permit the translation of overall goals into operational subgoals irrespective of the internal structure of the firm that accountants may be dealing with.

Therefore, just as we have not limited our attention to any particular goals, we shall not limit our attention to cases which involve a single goal only.

In this section we shall first approach to a problem of goal analyses which involve multiple goals by goal programming technique introduced in the previous chapter, and then focus our attention to specific cases in which multiple goals are incompatible.[2]

[2] C.f., e.g., Whinston, 1962. Also, see Cyert and March, 1963, Chapter VI, on sequential attention to multiple goals.

2. A Linear Programming Model of Multiple Goals

We now extend the models discussed in the previous chapter to represent cases of multiple goals.

Suppose that we have m goals whose levels are expressed by a combination of m real numbers $(b_1, b_2, \ldots, b_m)$ or by an m-component column vector b, and that these goals can be attained by linear combinations of n subgoal variables represented by an n-component column vector x. Let A be an m x n matrix which represents the relationship between goals and subgoals. Then, we can state the problem as follows:

$$(3.9) \qquad \begin{aligned} Ax &= b \\ x &\geq 0. \end{aligned}$$

This may be rewritten equivalently as follows assuming that a solution exists for (3.9).

$$(3.10) \qquad \begin{aligned} &\text{Minimize} && ey^+ + ey^- \\ &\text{subject to} && Ax - Iy^+ + Iy^- = b \\ & && x, y^+, y^- \geq 0 \end{aligned}$$

where e is an m-component row vector[1/] whose elements are all equal

---

1/ We shall discuss the commensurability of goals in the next subsection. Note that we need not be concerned with the commensurability of goals if the simultaneous attainment of goals is possible since then, $y^+ = y^- = 0$.

to 1, y^+ and y^- are m-component column vectors for the variables which were introduced in the previous chapter, I is an m-dimensional identity matrix, and A, x, and b are the same as defined above. Note that (3.9) and (3.10) are something different than merely separately considering m single goal problems formulated in (2.32) and (2.33), respectively, in that (3.10) permits interactions between the subgoal variables, in the n-rowed vector, x, which are to be considered simultaneously relative to all goals, b, in the minimization of

$$ey^+ + ey^- = \sum_{i=1}^{m} (y_i^+ + y_i^-).$$

<u>Example 3:</u> A production manager faces a problem of job allocation between his two teams which are identical with each other except for their processing rates, which are 1 unit an hour for Team 1 and .5 unit an hour for Team 2. The production manager wants to have the total daily production as close to 12 units as possible and also wants the daily work loads for each team be as close to 8 hours as possible. Let x_1 and x_2 be the working hours of Team 1 and Team 2, respectively. Then, the production manager's problem is formulated as follows:

$$(3.11)\quad \begin{array}{l} \text{Minimize} \quad y_1^+ + y_2^+ + y_3^+ + y_1^- + y_2^- + y_3^- \\ \text{subject to} \end{array}$$

$$\begin{array}{rcl} x_1 + .5x_2 - y_1^+ + y_1^- & = & 12 \\ x_1 - y_2^+ + y_2^- & = & 8 \\ x_2 - y_3^+ + y_3^- & = & 8 \\ x_1, x_2, y_1^+, y_2^+, y_3^+, y_1^-, y_2^-, y_3^- & \geq & 0 \end{array}$$

The solution to (3.11) is $x_1 = 8$, $x_2 = 8$, and all goals are attained with $y_1^+ + y_2^+ + y_3^+ + y_1^- + y_2^- + y_3^- = 0$.

3. Ordering and Weighting of Goals

In Example 3, we had a case in which all goals can be satisfied simultaneously. This is not always the case in real business problems. Suppose, for example, that the daily production level in Example 3 is raised to 15 units. Then, we can no longer have all goals satisfied by any combination of non-negative x_1 and x_2.

We now want to consider cases of incompatible multiple goals by reference to orderings and weightings of goals. That is, we first order these incompatible multiple goals so that goals in a lower rank are satisfied only after those in a higher rank are satisfied or have reached points beyond which no improvements are possible under the given constraints. We then give weights to goals in a same rank.

To see how this might be done "rationally" we can proceed as follows. First we check each one of m goals and ask whether over-(or under-) attainment of the goal is satisfactory. In the light of the answers received, we can then drop a y_i^+ (or y_i^-), where i is an index for the goal, from the functional. (See (2) and (3) in Table III-1.) If the goal is something which must be attained exactly (i.e., neither over nor under attainment of the goal is satisfactory), we leave both y_i^+ and y_i^- in the functional.

Alternatively, if a goal is to be minimized or maximized we change the signs of y_i^+ and y_i^- in the functional as given in (4) and (5) of Table III-1.

Secondly, we order the y_i^+ and y_i^- variables $(i = 1, 2, \ldots, m)$ which remain in the functional, starting from a positive or a negative deviation from a goal which is least important and continue the rankings until one is achieved which is "most important." Suppose we have obtained, in this way, goals classified in k ranks. We then assign to each variable in the j^{th} rank $(j = 1, 2, \ldots, k)$ a so-called "preemptive priority factor" M_j, $(j = 1, 2, \ldots, k)$. This preemptive priority is interpreted via the symbol $M_{j+1} >>> M_j$ $(j = 1, 2, \ldots, k-1)$ to mean that no number k, however large, can make kM_j greater than or equal to M_{j+1}.1/

1/ Charnes and Cooper, 1961, pp. 756-57, for a discussion of this so-called "non-archimedian order property."

In ordering the variables, such as y_i^+ and y_i^-, that will measure deviations from goals, we may encounter cases where we need finer divisions. Note our approach is one that first ascertains goals which have already been achieved. The remaining y_i^+ and $y_i^- \neq 0$ can then be placed in exact correspondence with the remaining unachieved goals by virtue of the condition $y_i^+ \cdot y_i^- = 0$, as already discussed. On the other hand, the ordering to be undertaken may depend upon the distances associated with the respective y_i^+, $y_i^- > 0$ which remain.

Such a dependence of the ordering upon the distance can be approximated by a step function characterized by a few "critical points"

or pseudo goal levels, just like the cases we discussed in Section 2 of Chapter II. Take, for example, $y_i^+ > 0$ which shows the degree of over-achievement of the i^{th} goal over and above its goal level b_j. Suppose we divide the range for y_i^+ into $(s + 1)$ segments by introducing s pseudo goal levels, $b_{i1}^+, b_{i2}^+, \ldots, b_{is}^+$, where $b_i =$ $b_{i0}^+ < b_{i1}^+ < b_{i2}^+ < \ldots < b_{is}^+$. Then, define variables y_{ij}^+ $(j = 0, 1, \ldots, s)$ as follows:

$$
(3.12)\quad
\begin{aligned}
y_{ij}^+ &= \text{Max}\,[\text{Min}\,(b_{i(j+1)}^+ - b_{ij}^+,\ y_i^+ - b_{ij}^+),\ 0]\quad (j = 0, 1, \ldots, s-1)\\
y_{is}^+ &= \text{Max}\,[\,y_i^+ - b_{is}^+,\ 0],
\end{aligned}
$$

so that, then,

$$
(3.13)\qquad y_i^+ = \sum_{j=0}^{s} y_{ij}^+ ,
$$

and the one goal deviation measure y_i^+ is decomposed into a sum of goal measures $y_{ij}^+ \geq 0$, some of which may be achieved even when $y_i^+ > 0$ obtains.

For the variables we may further assign different preemptive priority factors $M_{v_{ij}}$ $(j = 0, 1, \ldots, s)$ provided $M_{v_{i(j+1)}} >>> M_{v_{ij}}$. Then, by attaching the following set of equations to (3.10) we can obtain the desired property for y_{ij}^+ as given in (3.12).

$$
(3.14)\qquad y_i^+ - \sum_{\ell=j}^{s} y_{i\ell}^+ + \sum_{\ell=j}^{s} y_{i\ell}^- = b_{ij}^+ - b_i \quad (j = 0, 1, \ldots, s)
$$

Similar decomposition on y_i^- can be made as follows. First, set up s pseudo goal levels, b_{i1}^-, b_{i2}^-, ..., b_{is}^-, where $b_i = b_{i0}^- > .. > b_{is}^-$, and define variables y_{ij}^- ($j = 0, 1, \ldots, s$) as;

$$(3.15) \quad \begin{aligned} y_{ij}^- &= \text{Max}\,[\min\,(b_{ij}^- - b_{i(j+1)}^-,\ b_{ij}^- - y_i^-),\ 0]\ (j=0, 1, \ldots, s-1) \\ y_{is}^- &= \text{Max}\,[b_{is}^- - y_i^-,\ 0] \end{aligned}$$

Then, we have

$$(3.16) \quad y_i^- = \sum_{j=0}^{s} y_{ij}^-.$$

The property (3.15) for y_{ij}^- ($j=0, 1, \ldots, s$) may be obtained by attaching the following set of equations to (3.10).

$$(3.17) \quad -y_i^- - \sum_{\ell=j}^{s} y_{i\ell}^+ + \sum_{\ell=j}^{s} y_{i\ell}^- = b_{ij}^- - b_i \ (j=0, 1, \ldots, s)$$

Finally, those slack variables[1/] which are in a same order

1/ These are actually "natural slack" variables. See Charnes and Cooper, 1961. However, we shall simply call them slack variables unless the distinction becomes important.

---

group (i.e., a same coefficient M_i being attached) have to be weighted. The criterion is how much increase in a variable would be just offset by a unit decrease in some other variable in the

same order group. In other words, we want to minimize the sum of "regret"[1/] from all unsatisfactory achievement reflected in positive

---

1/ The term "regret" here is not used in precisely the same way as in economics, e.g., in game theory. See pp. 781 ff. in Volume II of Charnes and Cooper, 1961.

---

values of slack variables in the same order group, and the weighting factor α_i (which is a positive number) attached to the i^{th} slack variable in the functional represent a relative amount of regret on one unit of unsatisfactory deviation from the goal level.

In this sense, deviations from goals which are in the same order group must be commensurable. However, we can postpone such a weighting by giving an equal weight to each variable in the same order group, and then solve the problem. If all these variables are zero in the solution, we do not need to be bothered by weighting. If, however, one of the variables in the group is positive, which implies that every variable in the group can be made positive, then we determine the weights by careful studies of the situation and by possibly further splits of the variables as described above.

Such a gradual refinement of the problem is applicable in ordering slack variables, too. We can first classify the variables into two or three groups, solve the problem, then reorder or assign weights to the variables which are positive in the solution, forgetting all variables which are zero in the solution.

<u>Example 4</u>: Consider the following modification of Example 3. The production manager wants <u>most</u> to avoid any underachievement on the production level which has now been raised to 15 units. If overtime operation is necessary to attain the given production level, he wants to assign it to Team 1 and Team 2 in such a way that (i) any overtime operation for Team 1 beyond 2 hours is avoided and, <u>then</u>, (ii) minimize the sum of "regret" from assigning overtime operations to Team 1 (1 unit of "regret" for each hour of overtime operation) and to Team 2 (2 units of "regret" for each hour of overtime operation). He has some desire to avoid any idle time for both teams (with an equal "regret" factor) but this is to be done only after everything mentioned above is satisfied. Finally, any over-achievement in the production level is satisfactory for him but he has no particular interest in the production level once it passes the target.

The production manager's goals are now formulated as follows:

(3.18)

$$\text{Minimize} \quad M_2 y_2^+ + 2M_2 y_3^+ + M_3 y_{21}^+ + M_4 y_1^- + M_1 y_2^- + M_1 y_3^-$$

subject to

$$\begin{array}{lllllllll} x_1 + .5x_2 - y_1^+ & & & & + y_1^- & & & & = 15 \\ x_1 & - y_2^+ & & & & + y_2^- & & & = 8 \\ \quad x_2 & & - y_3^+ & & & & + y_3^- & & = 8 \\ & y_2^+ & & - y_{21}^+ & & & & + y_{21}^- & = 2 \end{array}$$

Note that y_1^+ does not come into the functional since the manager does not have any interest in the production level once it passes the target. The fourth constraint is included so that any overtime operation for Team 1 beyond 2 hours (represented by y_{21}^+) is avoided with higher priority than other overtime operations.

The solution is $x_1 = 10$, $x_2 = 10$, with $y_2^+ = 2$, $y_3^+ = 2$, and all other variables equal to zero.

In general, coefficients in the functional of a goal programming problem consist of a combination of preemptive priority factors

M's for ordering and weighting factors α's. Let c be a 2m-component row vector whose elements are products of α and M_v, i.e.,

$$c = (\alpha_1 M_{v_1}, \alpha_2 M_{v_2}, \ldots, \alpha_{2m} M_{v_{2m}}) \tag{3.19}$$

where α_i's (i=1, 2, ..., 2m) are real numbers and M_{v_i} (i = 1, 2, ..., 2m: v_i = 1, 2, ..., k) are preemptive priority factors with the highest preemptive priority factor being M_k. Let y be a 2m-component column vector whose elements are y^+'s and y^-'s, i.e.,

$$y = [y_1^+, y_2^+, \ldots, y_m^+, y_1^-, y_2^-, \ldots, y_m^-]. \tag{3.20}$$

Then, a problem of goal analysis can be represented by the following general linear programming problem:

$$\begin{aligned} &\text{Minimize} && cy \\ &\text{subject to} && ax + Ry = b \\ & && x, y \geq 0 \end{aligned} \tag{3.21}$$

where A and R are matrices of the size m x n and m x 2m, respectively. This representation is general, of course, only on the understanding that the dimensions of c, y, and R are increased and structured appropriately when y_{ij}^+'s and y_{ij}^- are in an expanded y for (3.20).

Note the difference in the role of x and y in (3.21) compared with an ordinary linear programming formulation. In a goal programming model like the one formulated in (3.21), the slack variables, y, are

used to "drive" the structural variables, x, whereas in an ordinary linear programming problem it is the structural variables which are the driving force and the slack variables are adjusted accordingly.

In summary, what we have to do in order to derive a solution to an incompatible multiple goal problem, is to <u>order</u> and <u>weight</u> under- and over-achievement of each one of multiple goals in terms of preferences from the viewpoint of overall business operations. We have reviewed a technique which can approximate cases where these preferences are non-linear functions of the deviations from the goal levels.

Of course a set of constraints on subgoals in the form of $Bx \leq h$ can be directly incorporated in (3.21), in which case the model becomes

$$
\begin{aligned}
&\text{Minimize} && cy \\
(3.22)\qquad &\text{subject to} && Ax + Ry = b \\
& && Bx \leq h \\
& && x, y \leq 0
\end{aligned}
$$

Note that the constraints $Bx \leq h$ may be included in the matrix A in the form of $Bx - y_b^+ + y_b^- = h$ with <u>highest</u> preemptive priority factors attached to the variables y_b^+. The difference between the term "goals" and the term "constraints" is that the former represents the manager's desires, whereas the latter represents the environment

of his operation. However, in the mathematical formulation, the only difference between the two is that the constraints have to be satisfied before any attempt is made to meet the goals. In other words, constraints have a higher preemptive value than any of the goal preemptions incorporated in the functional, even when they cannot be satisfied under the circumstances of incompatible goals. Thus, the two can be incorporated together in the linear programming formulation (3.21) with only a difference in the preemptive priority factors. In any event the goals must be satisfied insofar as the constraints permit this to be attained. When it is not possible to satisfy all goals then, considering all interactions permitted by the constraints, or optimum is finally achieved only in accordance with the orderings and weightings of goals that are incorporated in the functional.

In the next section we shall approach the same problems of multiple goals by means of generalized inverses, noting the difference in the way in which deviations from the goal levels are minimized under the linear programming approach discussed in this section and the generalized inverse approach which is the topic for the next section.

4. Analysis of Multiple Goals: A Generalized Inverse Approach

1. A Generalized Inverse Approach

Following this goal programming approach to multiple goals we now try to secure further perspective and insight by approaching this same topic in a different way by using the generalized inverse of a matrix applied to the relationships between goals and subgoals.

Let b be an m-component column vector for goal levels and x an n-component column vector for subgoal variables which are related to the constants b by A, an m x n matrix of known constants. Then, the problem of multiple goals is to find a set of solution subgoals x which attain the multiple goals simultaneously, or to find x which satisfy,

$$Ax = b \tag{3.23}$$

As discussed in detail in Appendix A, any solution to (3.23), provided a solution exists, can be written as,

$$x = A^{\dagger}b + A^{o}z \quad (z \in E^{n-r}) \tag{3.24}$$

where $A^{\dagger}$, which is n x m, is the generalized inverse of A. The latter is assumed, without loss of generality, to have rank r, in which case A^{o} is an n x (n-r) matrix of a null space basis which has the property $AA^{o} = 0$, and z is an (n-r)-component arbitrary vector. By allowing z to vary over all of (n-r)-dimensional Euclidean space (E^{n-r}), we can exhaust all the solutions to $Ax = b$ by (3.24). (Refer to Section 5 of Appendix A.)

The first term in (3.24) shows a <u>particular</u> solution to $Ax = b$. This particular solution has certain characteristics which will be discussed in detail later. For the moment we need only observe that when $Ax = b$ has a solution then we can derive the particular solution from the general one in (3.24) by merely choosing $z = 0$ to obtain

$$x = A^{\dagger}b. \tag{3.25}$$

The discussions and a proof that $x = A^{\dagger}b$ is a solution to $Ax = b$ (provided a solution exists) are in Section 1 - 5 of Appendix A.

While (3.25) is a solution to $Ax = b$, it is not, in general, the only one. To see this we can proceed as follows: Take any vector x^o in the null space of A, denoted by $N(A)$, i.e., any vector x^o such that $Ax^o = 0$, and add it to a solution $x^*(= A^{\dagger}b)$. Then, clearly,

$$A(x^* + x^o) = Ax^* + Ax^o = b + 0 = b. \tag{3.26}$$

This means that if x^* is a solution to $Ax = b$ then the vector $(x^* + x^o)$ is also a solution for any x^o vector in $N(A)$, the null space of A. On the other hand, $N(A)$ is a linear vector space. Since every $x^o \in N(A)$ may be uniquely represented by expressions of the form $A^oz = x^o$ where z is an $(n-r)$-component column vector, in which event, we have $Ax^* + AA^oz = Ax^* + 0 = b$ in place of (3.26).

Given any A^o, a basis for N(A), we have only a unique z for each x^o in $A^o z = x^o$. But this does not mean that A^o is unique since the many different bases may be specified for N(A). On the other hand, when uniqueness is wanted, as is often the case, it may be attained by using instead of A^o the n x n matrix $(I - A^\dagger A)$ which is unique since (1) the identity matrix I is unique, and (2) $A^\dagger$ is unique so that $A^\dagger A$ is also unique. Also every $x^o \varepsilon N(A)$ may also be represented as $(I-A^\dagger A)z=x^o$ so that nothing is lost when this expression is used in lieu of $A^o z = x^o$ for $x^o \varepsilon N(A)$.

As observed in Appendix A, a vector $(I - A^\dagger A)z$ is associated with any given z via a least sum of squares--or a minimum distance in the sense of a Euclidean metric. It is also interesting, however, to pursue our goal analysis here by means of z choices in order to see how these may be related to managerial discretion in the choice of subgoal values. Any constraints which we may want to impose upon subgoal variables x, including the non-negativity condition, may be considered as constraints on this discretionary area. For example, if x is constrained to be non-negative, then we can no longer take any vector x for an A^o basis of N(A). Instead, we must restrict choices of the weighting factors in z so that they will satisfy

$$A^\dagger b + A^o z \geq 0, \text{ or} \tag{3.27}$$

$$A^o z \geq -A^\dagger b.$$

If we further want to impose $Bx \leq h$, then the discretionary choices of z must be further confined by

$$BA^{\dagger}b + BA^{o}z \leq h \quad \text{or}$$

$$BA^{o}z \leq h - BA^{\dagger}b$$

in order to ensure that (3.24) will then give x values within the discretionary area that is allowed.

Example 5: Take Example 1 in Section 2 of this chapter. Let us consider the limit on the machine operating time in the example not as constraints but as the manager's desire or goals. Then, this single goal problem with constraints on subgoals may be interpreted as a multiple goal problem which involves three goals: (i) attain the profit level of \$1 after covering the fixed costs of \$1.5, (ii) do not operate Machine 1 beyond 12 hours, and (iii) do not operate Machine 2 beyond 10 hours.

Let x_1 and x_2 be the sales (and production volume) of Product 1 and Product 2, and let x_3 and x_4 be the idle time of Machine 1 and Machine 2, respectively. Then, this problem is represented by the following set of equations.

$$
(3.29)\quad
\begin{aligned}
x_1 + .5x_2 \quad\quad\quad &= 2.5\\
3x_1 + 2x_2 + x_3 \quad &= 12\\
5x_1 \quad\quad\quad + x_4 &= 10\\
x_1,\ x_2,\ x_3,\ x_4 &\geq 0
\end{aligned}
$$

This is evidently of the form $Ax = b$, where

$$
A = \begin{bmatrix} 1 & .5 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 5 & 0 & 0 & 1 \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \qquad \text{and } b = \begin{bmatrix} 2.5 \\ 12 \\ 10 \end{bmatrix}
$$

(Compare this formulation with (3.3).) The generalized inverse of A is

$$
A^{\dagger} = \frac{1}{62}\begin{bmatrix} 16 & -2 & 10 \\ 92 & 4 & -20 \\ -232 & 60 & 10 \\ -80 & 10 & 12 \end{bmatrix}
$$

and a possible null space basis is given by[1/]

$$
A^{o} = \frac{1}{62}\begin{bmatrix} 2 \\ -4 \\ 2 \\ -10 \end{bmatrix}
$$

[1/] Refer to Section 4 and 6 of Appendix A for methods of calculating A^{o} and $A^{\dagger}$.

Therefore, the set of all solutions to (3.29) for any goal level, b, provided, for the choice of b, a solution exists for Ax = b is given by :

(3.30)
$$x = \frac{1}{62}\begin{bmatrix} 16 & -2 & 10 \\ 92 & 4 & -20 \\ -232 & 60 & 10 \\ -80 & 10 & 12 \end{bmatrix} b + \frac{1}{62}\begin{bmatrix} 2 \\ -4 \\ 2 \\ -10 \end{bmatrix} z,$$
$$x \geq 0,$$

where z is an arbitrary scalar. (Cf. (3.24).) For (3.29) we have

$$b = \begin{bmatrix} 2.5 \\ 12 \\ 10 \end{bmatrix}$$

(3.31)
$$x = \frac{1}{62}\begin{bmatrix} 116 \\ 78 \\ 240 \\ 40 \end{bmatrix} + \frac{1}{62}\begin{bmatrix} 2 \\ -4 \\ 2 \\ -10 \end{bmatrix} z$$

But now by virtue of the non-negativity on x we cannot choose z in a wholly arbitrary manner.

The non-negativity condition on x is transformed into the constraint on z which is calculated as

(3.32) $\quad -58 \leq z \leq 4.$

Thus, the set of all admissible solutions to (3.29) is given by (3.31) and (3.32) together.

We are also in a position to study the subgoal consequences of altering any goal. For instance if the profit level is raised to \$2 over and above the fixed costs of \$1.5, we change the goal vector b to $\begin{bmatrix} 3.5 \\ 12 \\ 10 \end{bmatrix}$, and put this into (3.30), obtaining,

$$(3.33) \qquad x = \frac{1}{62}\begin{bmatrix}132\\170\\8\\-40\end{bmatrix} + \frac{1}{62}\begin{bmatrix}2\\-4\\2\\-10\end{bmatrix} z,$$

$$x \geq 0$$

Note that in this case the non-negativity condition of x is satisfied if and only if $z = -4$. See the last <u>two</u> lines of (3.33). Therefore, even though there are infinitely many solutions to (3.29) with the goal vector

$$\begin{bmatrix}3.5\\12\\10\end{bmatrix}$$

, we have only one solution if we impose the non-negativity condition on x. The unique solution is

$$x = \begin{bmatrix}2\\3\\0\\0\end{bmatrix}$$

, and altering the first goal from $b = 2.5$ to $b = 3.5$ means that no discretion is to be allowed for any of the subgoal variables even though the x_3 and x_4 of (3.29) do not enter directly in the "profit goal" constraint. For the level of of the first goal higher than 3.5, there is no value of z which can satisfy the non-negativity condition of x. Hence, if $z > 3.5$ is prescribed, an incompatible multiple goal situation is encountered.

2. <u>Incompatible Multiple Goals</u>

So far, we have analyzed the multiple goal problem in terms of A^{o} representations only on the assumption that there is a set of values for subgoals by which we can attain all of the multiple goals simultaneously, i.e., assuming that solution exists for $Ax = b$. Suppose that there is no way of attaining all of the multiple goals simultaneously, i.e., no solution exists for $Ax = b$.

In such cases, (3.24) gives us solutions which come "closest" to the simultaneous attainment of the multiple goals due to the "least square property" of $A^{\dagger}$, the generalized inverse, which is discussed in Section 5 of Appendix A. Here, the degree of "closeness" is measured by the square root of the sum of the squares of the deviations from the goals -- i.e., a Euclidean or ℓ_2 metric -- whereas in a goal programming approach we were using a so-called ℓ_1 metric.[1]

1/ C.F., Charnes and Cooper, 1961, Appendix A. In general, an ℓ_p metric, or measure of distance, for any finite dimensional vector space is given by

$$d_p = \left(\sum_{i=1}^{n} | x_i - y_i |^p \right)^{1/p}$$

where x_i and y_i are the respective components of two points x and y.

To bring the points at issue here together we repeat the previous section, i.e.,

$$\begin{array}{lll} & \text{Minimize} & ey^+ + ey^- \\ (3.10) & \text{subject to} & ax - Iy^+ + Iy^- = b \\ & & x, y^+, y^- \geq 0 \end{array}$$

and observe that this minimization also solves

$$(3.34) \qquad \text{Minimize} \sum_{i=1}^{m} |A_i x - b_i|$$

where A_i is the i^{th} row of A and b_i is the i^{th} element in b, since evidently $y_i^+ \geq 0$ $y_i^- \geq 0$ gives $y_i^+ + y_i^- = |A_i x - b_i| \geq 0$ with $y_i^+ \cdot y_i^- = 0$ in the minimization. On the other hand, solutions obtained from (3.24) by the use of the generalized inverse of A is equivalent to

$$(3.35) \qquad \text{Minimize } [\sum_{i=1}^{m} (A_i x - b_i)^2]^{1/2}$$

Thus, if we consider the ℓ_p metric to be given by,

$$(\sum_{i=1}^{m} |A_i x - b_i|^p)^{1/p}$$

then (3.34) is applicable for p = 1 and (3.35) for p = 2.

This means that when multiple goals are incompatible, in the sense that they cannot all be satisfied simultaneously, then both the goal programming and generalized inverse approaches derive solutions which minimize the "distance" from the set of goals. But distance in the goal programming approach is measured by the sum of absolute value of the deviations from each goal whereas the distance in the generalized inverse approach is measured by the square root of the sum of squares of the deviation from each goal. Of course the resulting x values will differ as we prescribe different ℓ_p measures such as ℓ_1 or ℓ_2 for our minimization. This too represents a problem of managerial choice in goal programming.

We shall give the following example to show the minimization of an ℓ_2 metric.

Example 6: Suppose that we change the goals in Example 5 as follows: (i) attain the profit level of \$2.5 over and above the fixed costs of \$1.5, (ii) operate Machine 1 for 12 hours, and (iii) operate Machine 2 for 10 hours. Then, no choice of x can satisfy all of the above three goals, since there is no solution to

$$(3.36) \qquad \begin{aligned} x_1 + .5x_2 &= 4 \\ 3x_1 + 2x_2 &= 12 \\ 5x_1 &= 10 \\ x_1, x_2 &\geq 0. \end{aligned}$$

The set of equation in (3.36) may be represented by Ax = b where

$$A = \begin{bmatrix} 1 & .5 \\ 3 & 2 \\ 5 & 0 \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad b = \begin{bmatrix} 4 \\ 12 \\ 10 \end{bmatrix}$$

The generalized inverse of A is

$$A^\dagger = \frac{1}{426}\begin{pmatrix} 4 & -1 & 85 \\ 44 & 202 & -130 \end{pmatrix}$$

and the null space of A consists of only the zero vector, as can easily be verified. Therefore, by (3.24) we have

$$(3.37) \qquad x = \frac{1}{426}\begin{pmatrix} 4 & -1 & 85 \\ 44 & 202 & -130 \end{pmatrix}\begin{pmatrix} 4 \\ 12 \\ 10 \end{pmatrix} = \begin{pmatrix} 2.005 \\ 3.052 \end{pmatrix}.$$

However, Ax is then, $\begin{pmatrix} 1 & .5 \\ 3 & 2 \\ 5 & 0 \end{pmatrix}\begin{pmatrix} 2.005 \\ 3.052 \end{pmatrix} = \begin{pmatrix} 3.53 \\ 12.12 \\ 10.02 \end{pmatrix}$,

which is not equal to $\begin{bmatrix} 4 \\ 12 \\ 10 \end{bmatrix}$. Thus, the simultaneous attainment of the goal is impossible and a solution which comes closest to the simultaneous goal attainment is given by (3.37) — closeness being measured by the square root of the sum of the deviations from each goal, i.e.,

$$(3.38) \qquad d = [(3.53 - 4)^2 + (12.12 - 12)^2 + (10.02 - 10)^2]^{1/2}$$

so that $x_1 = 2.005$, $x_2 = 3.052$, as in (3.37), is minimum, in the least squares sense, among all possible subgoal values. (Note that the non-negativity condition of x is satisfied in this example without any constraint on the "discretionary" area since we have only the null vector at our discretion in this case.)

We now proceed to check our solutions $x_1 = 2.005$ and $x_2 = 3.052$, by reference to the ordinary route of a least squares minimization. Referring to (3.36) we form the square root of sum of squares to be minimized, or, equivalently, in this case, the sum of squares to be minimized as

$$\phi(x_1, x_2) = (x_1 + .5\ x_2 - 4)^2 + (3\ x_1 + 2\ x_2 - 12)^2 + (5\ x_1 - 10)^2 .$$

Differentiating first on x_1 and then on x_2 and equating to zero produces the so-called normal equations for a least squares solution[1/]--

1/ This example has been devised so that we can ignore things like the Kuhn-Tucker conditions relative to the non-negativity requirements on x_1 and x_2. Cf. H. W. Kuhn and A. W. Tucker, 1951.

viz.,

$$70\ x_1 + 13\ x_2 = 180$$
$$13\ x_1 + 8.5\ x_2 = 52$$

Solving for x_1 and x_2 produces $x_1 = 2.005$ and $x_2 = 3.052$ which agrees with our previously attained results.

Minimization of an ℓ_1 or an ℓ_2 metric as an ordinary scalar implies that the multiple goals must be commensurable. This is not always wanted and so we shall therefore proceed to an analysis of goal orderings where commensurability is not present. We shall do this, moreover, in a way that produces commensurability in the special kinds of goal weightings that have this property.

3. Ordering and Weighting of Goals

In order to solve incompatible multiple goal problems, we proceed as follows.

Suppose $Ax = b$ is partitioned into $A_i x = b_i$ $(i = 1, 2, \ldots, k)$ in accordance with an ordering of the goals b. We assume that b_k is a subvector of b and contains the goals which have the highest priority, b_{k-1} is a subvector for goals with the second highest priority,...., and finally b_1 is a subvector for goals with the lowest priority. Also we use A_i for a matrix prepared by adjoining a set of row vectors of A with rows that correspond to a subvector b_i.

Now, then, the set of all subgoals which satisfy the goals with the highest priority is given by

$$x = A^{\dagger}_k b_k + A^{o}_k z_k \quad (z_k \in E^{n-r_k}) \tag{3.39}$$

where r_k is the rank of A_k. If $A_k x = b_k$ has a solution, (or, equivalently, if b_k is in the range of A_k), the set of all solutions to $A_k x = b_k$ is given by (3.39). If no solution exists for $A_k x = b_k$, (or, equivalently, if b_k is not in the range of A_k), then (3.39) gives us a set of solutions which come "closest" to the set of goals with highest priority, "closeness" being measured by the ℓ_2 metric.

We then move to the goals with the second highest priority, and, if possible, find the x which satisfy

$$A_{k-1} x = b_{k-1}. \tag{3.40}$$

However, x's which satisfy this equation have to be chosen from those represented by (3.39). Therefore, the restriction on x by (3.40) is transformed into the restrictions on the free variable z_k in accordance with:

$$A_{k-1}x = A_{k-1}A_k^{\dagger}b_k + A_{k-1}A_k^{o}z_k = b_{k-1} \quad \text{or} \tag{3.41}$$

$$A_{k-1}A_k^{o}z_k = b_{k-1} - A_{k-1}A_k^{\dagger}b_k \tag{3.42}$$

Note that although the right hand sided (3.32) is constant it does not follow that z_k is fixed by this expression. It only means that z_k must be chosen to satisfy

$$z_k = (A_{k-1}A_k^{o})^{\dagger}(b_{k-1} - A_{k-1}A_k^{\dagger}b) + (A_{k-1}A_k^{o})^{o}z_{k-1} \tag{3.43}$$

By substituting this last expression in (3.39), we obtain an expression for a set of x's which will minimize the distance from the goal levels with the highest priority. Then, subject to no worsening of the previous minimization, it will minimize the distance from the goal levels with the second highest priority. With this assurance then we can relate x to z_{k-1}, instead of z_k, by inserting (3.43) into (3.39) in order to obtain

$$x = A_k^{\dagger} b_k + A_k^{o}(A_{k-1} A_k^{o})^{\dagger}(b_{k-1} - A_{k-1} A_k^{\dagger} b) \tag{3.44}$$

$$+ A_k^{o}(A_{k-1} A_k^{o})^{o} z_{k-1} \qquad (z_{k-1} \in E^{n-r_{k,k-1}})$$

where $E^{n-r_{k,k-1}}$ is the Euclidean space of $n-r_{k,k-1}$ dimensions (n being the number of columns in A and $r_{k,k-1}$ being the rank of $\begin{bmatrix} A_k \\ A_{k-1} \end{bmatrix}$.)

We proceed in this fashion changing the free variables in the expression for x -- i.e., the z's -- into variables with fewer and fewer dimensions until the process is concluded at some stage with some $z_{k-\ell}$ fixed uniquely by its preceding expression or until all the priority orderings are considered at the stage to which z_1 and hence b_1 applies.

Suppose next that we desire to weight as well as order our goals.1/ This may be done in the following way. Suppose that we

---

1/ This may also be done in the goal programming formulation by means of the expressions

$$\sum_{i=1}^{m} \alpha_i \left| \sum_{j=1}^{n} a_{ij} x_j - b_i \right|$$

with all $\alpha_i \geq 0$.

want to measure the degree of "regret" by the square of the deviation from the goal level, and that this degree of "regret" is commensurable among all s goals in the same ordered rank by means of non-negative weighting factors α_j $(j = 1, 2, \ldots, s)$. Then, we want to minimize for goals in the same ordered rank, the total "regret" measure, d^2,

$$(3.45) \qquad d^2 = \sum_{j=1}^{s} \alpha_j (A_j x - b_j)^2 = \sum_{j=1}^{s} (\sqrt{\alpha_j}\, A_j x - \sqrt{\alpha_j}\, b_j)^2$$

If instead of obtaining d^2 directly as in (3.45) we proceed by means of generalized inverses we can achieve the same result if we multiply each element in A_j and b_j by $\sqrt{\alpha_j}$ $(j = 1, 2, \ldots, s)$. The least square property of the generalized inverse assures that the total "regret" measure, d^2, (or equivalently its square-root, d) is minimized by the resulting solution.

Example 7: Suppose that we modify Example 6 so that Product 2 does not require the operations of Machine 1 or Machine 2. Then, the problem (3.36) in Example 6 is changed to

$$(3.46)\qquad \begin{aligned} x_1 + .5x_2 &= 4 \\ 3x_1 &= 12 \\ 5x_1 &= 10 \\ x_1, x_2 &\geq 0. \end{aligned}$$

Suppose further that management wants to give the first goal (profit level) the highest priority of being satisfied. Then, the set of all x's which satisfy $x_1 + .5x_2 = 4$ is given by

$$(3.47)\qquad x = \begin{pmatrix} .8 \\ .4 \end{pmatrix} 4 + \begin{pmatrix} -.5 \\ 1 \end{pmatrix} z = \begin{pmatrix} 3.2 \\ 1.6 \end{pmatrix} + \begin{pmatrix} -.5 \\ 1 \end{pmatrix} z$$

where z is a scalar. (Refer to (2.25) in Chapter II.) Due to the non-negativity condition imposed upon x, z must be in the range of $-1.6 \leq z \leq 6.4$. (Refer to (2.29) in Chapter II.)

The first goal being satisfied, management then wants to satisfy the second and the thrid goals concerning the operating times of Machine 1 and Machine 2. Suppose that the manager is so concerned with the operating time of Machine 1 that he estimates his "regret" will be four times as much for any deviation in the Machine 1 operating time from its goal level as the same amount of deviation in the Machine 2 operating time from its goal level. In this case only proportionality in the weights matters. Hence we multiply each element in the second equation of (3.46) by 2, which is the square root of 4, while leaving each element in the third equation intact.

Now we want to solve,

$$(3.48)\qquad \begin{pmatrix} 6 & 0 \\ 5 & 0 \end{pmatrix} x = \begin{pmatrix} 24 \\ 10 \end{pmatrix}$$

which is derived from the second and the third equations of (3.46) after being adjusted for the weighting factor. However, the variable x must satisfy (3.47). Therefore, we replace x in (3.48) by the expression in (3.47), obtaining,

$$(3.49)\qquad \begin{pmatrix} 6 & 0 \\ 5 & 0 \end{pmatrix} \left[\begin{pmatrix} 3.2 \\ 1.6 \end{pmatrix} + \begin{pmatrix} -.5 \\ 1 \end{pmatrix} z\right] = \begin{pmatrix} 24 \\ 10 \end{pmatrix} \quad \text{or}$$

(3.50) $\quad \binom{-3}{-2.5} z = \binom{4.8}{-6}$.

By using the generalized inverse of $\binom{-3}{-2.5}$ which is $\frac{1}{15.25}$ $(-3. -2.5)$, (see page II-32a), we obtain,

(3.51) $\quad z = \frac{1}{15.25}(-3, -2.5)\binom{4.8}{-6} = \frac{.6}{15.25} = \frac{12}{305}$

Note that z is in the range specified above and, therefore, we can proceed without worrying about the non-negativity condition on x. By substituting this value of z into (3.37) we obtain $x_1 = \frac{194}{61}$, $x_2 = \frac{100}{61}$.

Note that the value of z derived in (3.51) does not satisfy (3.50). This means that the second and the third goals cannot be satisfied by any solution which satisfies the first goal, and that the solution which minimizes the total amount of "regret" defined above, i.e., the solution which minimizes

(3.52) $\quad d^2 = 4(3x_1 - 12)^2 + (5x_1 - 10)^2$

will satisfy $x_1 + .5x_2 = 4$. But it will not satisfy either $3x_1 = 12$ or $5x_1 = 10$, missing the former by $12 - 3x_1 = 3$ units and the latter by $10 - 5x_1 = 6$ units, approximately.

In this manner, we see that the ordering and weighting of goals can also be accomplished over an ℓ_2 metric by means of a generalized inverse approach to goal analysis. This generalized inverse approach is efficient for attacking the problems of multiple goals if the variables involved in the problem are not required to be non-negative, or, as in the above examples, if the non-negativity constraints are redundant. If, however, the problem involves variables which are constrained to be non-negative and if the non-negativity

constraints are critical in the solutions, then it may be better to resort to some other means such as quadratic programming for achieving an ℓ_2 minimization.

The same things can be said when the problem involves under- or over-achievement of goals. Clearly if we add a goal level (and its linear function) to the goals (and their linear functions) which have already been imposed upon a set of subgoals, the number of dimensions of the free variables z in the expression for subgoals x in (3.39) is reduced by one unless the newly added linear function can be derived by a linear combination of other linear functions which are already available. (The rank of the null space of A decreases by one each time we add a new row which is independent of the old rows.) However, in the generalized inverse approach, a set of inequalities given by $Ax \leq b$ needs to be converted into $Ax + Iy = b$ or

$$(3.53) \qquad [A, I] \binom{x}{y} = b$$

where y is a column vector which is constrained to be non-negative. Similarly, when we add a new goal and its linear function in the for of $A_i x \leq b_i$, we have to change it to an equation by adding a non-negative slack variable y_i, i.e., $A_i x + Iy_i = b_i$. This means that the dimension of the null space of [A, I] is equal to the number of subgoals x (or the number of columns of A) regardless of the number of goals (or the number of rows of A.)

Therefore, an application of the generalized inverse of [A, I] does not help to reduce the dimension of the free variable in (3.39) as is shown by means of the following.

$$(3.54) \qquad \binom{x}{y} = [A, I_m]^{\dagger} b + \begin{bmatrix} I_n \\ -A \end{bmatrix} z \qquad (z \in E^n)$$

$$y \geq 0$$

where A is an m x n matrix, I_m and I_n are identity matrices of dimension m and n, respectively, and the second component is an expression for the null space of [A, I]. Note that $[A, I_m]^{\dagger}$ is of order $(n + m) \times m$ when A is $m \times n$, and then let G and H be the matrices formed from the first n rows and the next m rows, respectively, of $[A, I_m]^{\dagger}$. Then (3.54) is equivalent to

$$(3.55) \qquad x = Gb + I_n z$$

$$Az \leq Hb$$

where the last expression comes from the non-negativity requirement on y -- viz., $Hb - Az \geq 0$, in (3.54). This means that original inequalities in $Ax \leq b$ are transformed into the inequalities $Az \leq Hb$ by the adjunction of slack to obtain (3.53) as an equivalent formulation. Of course, the successive application of $[A, I_m]^{\dagger}$ to $Az \leq Hb$, $Aw \leq H^2b$, ... does not simplify the problem. On the other hand, the properties of generalized inverses have yet to be exploited extensively for management problems[1/], and the discussion here and in Appendix A have been developed

1/ See, e.g., Charnes, Cooper and Thompson, December, 1962, for an interesting exception.

and introduced with this in mind.[2/]

2/ They have been used in engineering, physics, statistics and elsewhere. See, for example, Duffin, 1962; Penrose, 1956, Zelen, 1962, etc.

CHAPTER IV

FEEDBACK ARRANGEMENTS FOR INFORMATION ON GOAL ATTAINMENTS

1. Introduction

In Chapter II and III, we explored the possibilities of applying linear programming technique and generalized inverse to the analysis of goals, a problem which was formulated as an extension of breakeven analysis. In those chapters, the discussion was directed to the problem: how to decompose a set of given goals into more operational subgoals. There it was assumed that subgoals could be more easily controlled by management than the goals which a management wants to attain. Now we introduce and examine one of the pertinent properties which, at least for accounting purposes, is closely related to control and in some respects is a sine qua non for accounting treatment.[1] This property is "measurability for control," -- or, alternatively, "control measurability." By "control measurability" we refer to the degree of ease or difficulty in obtaining feedback concerning actual goal attainment in practical

[1] At least of the traditional and the current versions of accounting. C.f., e.g., Trueblood, 1950. Possible future developments like the inclusion of "personality profiles," etc. in controllership records, etc., are not at issue here. C.f., e.g., Churchill, 1962 or Stedry, 1960.

situations. As a very simple example, the net profit from operations has less measurability for management control than the total dollar sales since the former depends upon the latter plus other factors such as costs. Moreover, as this simple example already suggests, we do not want to confine ourselves to the rules of ordinary arithmetic but we also include relations of "order" and indicators with more or less information content in the kinds of control measures we shall consider. Thus the sales figure, for instance, admits of sharper interpretations than does net profit insofar as the latter includes items like depreciation, etc., where precise arithmetic is not always accompained by equally precise interpretations.

In goal analyses developed in the previous chapters, we required subgoals to be more operational (controllable) than the goals to which they might be related, but we did not require subgoals to have more "control measurability." In this manner we emphasize that operationality is always essential for any rational scheme of subgoal formation. The supposition is then that we can always devise a system of measures, even for so called qualitative subgoals, although we cannot of course restrict ourselves to the narrow confines of ordinary arithmetic for this purpose and we may have to content ourselves with information loss as well. In any event we find it generally more reasonable to proceed from operationality to measurability as a criterion rather than the reverse.

2. The Relationship Among Goals, Subgoals, and Indicators

Given a separation of operationality and measurability, as suggested in the previous section, it is now natural to examine possible kinds of "indicators" in order to ascertain their "control measurability" properties relative to the goals and the subgoals in question. This in turn will necessitate our examining the relationships that might obtain between indicators and subgoals with special reference to their use in measuring the degree of goal attainment.

The following mathematics will be helpful in clarifying the problem: Let b be an n-component column vector for goal parameters which are the given targets for a set of goal variables and let v be an n-component column vector for goal variables. The distinction between goal parameters and goal variables is made here to keep track of differences that might arise at various stages -- e.g. in a planning stage when goals are set equal to a set of constants (e.g., aspiration levels)[1] and then subsequently, in a

[1] Including "aspirations about aspirations" as reflected in the budgeting -- control processes discussed in Stedry, 1960.

control stage when actual values of goal variables (e.g. actual cost) are compared with the given goal parameters(e.g., a standard or budgeted cost.)[1]

[1] We may also want to distingush different goal parameters, too, by according one of them the status of a "constant" and the other the status of a "variable."

As distinguished from b and v we now introduce the symbol w for a k-component column vector whose entries are "indicators" that are supposed, for one reason or another, to be more easily measured in a practical situation than either goals or subgoals. We assume, as in the previous chapters, that the relationship between the goal variables v and the subgoal variables x is given by,

$$(4.1) \qquad Ax = v$$

where A is an m x n matrix with entries representing technological considerations, legal or policy considerations, etc.

We now want to measure the value of each element in v, i.e., the value of each one of m goals, from time to time, by means of k indicators, $w_1, w_2, \ldots w_k$, whose values are determined by a linear combination of n subgoal variables, $x_1, x_2, \ldots, x_n$, i.e.

$$(4.2) \qquad Cx = w$$

where C is a k x n matrix and x and w are n- and k-component column vectors, respectively.

The characteristics of each indicator are now completely described by the corresponding row of C. It is known that there is not, in general, a consolidation of A into a smaller C which preserves all of the original information relative to v and x. 1/ This is not

1/ An interesting development of this for the so-called Leontief matrices of input-output analysis is given in Appendix E of Charnes and Cooper, 1961, which can be related to the goal programming concepts that were discussed earlier in this thesis.

what we have in mind, however. In order to show what we intend to represent by the matrix C, consider the following example.

<u>Example 1</u>: Suppose a firm had three sales transactions in a particular day, each one of which involves sales of Product 1 and Product 2. Let x_1, x_2, and x_3 represent the sales quantities of Product 1 in the first, the second, and the third transactions, respectively, and let x_4, x_5, and x_6 represent the sales quantity of Product 2 in the first, the second, and the third transactions, respectively. Also, let p_1, p_2, and p_3 be the selling prices of Product 1 applied to the three respective transactions and p_4, p_5, and p_6 be the selling prices of Product 2 applied to the three respective transactions.

Then the following is an example of the a set of 6 indicators that are produced from 6 subgoal Variables denoted by a 6-component column vector x.

$$Cx = \begin{bmatrix} p_1, & p_2, & p_3, & p_4, & p_5, & p_6 \\ p_1, & p_2, & p_3, & 0, & 0, & 0 \\ 0, & 0, & 0, & 1, & 1, & 1 \\ p_1, & 0, & 0, & p_4, & 0, & 0 \\ 1, & 0, & 0, & 0, & 0, & 0 \\ 0, & p_2, & 0, & 0, & 0, & p_4 \end{bmatrix} x = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \end{bmatrix}$$

The first indicator, w_1, shows the total dollar sales in the day. The second indicators, w_2, shows the total dollar sales of Product 1 in the day, while w_3 shows the total physical sales quantity of Product 2 in the day. The forth indicator, w_4, shows the total amount of sales in the first transaction, while w_5 shows specifically the value of x_1, the physical sale quantity of Product 1 in the first transaction. The sixth indicator, w_6, is the sum of the dollar sales of Product 1 in the second transaction and the dollar sales of Product 2 in the third transaction. Note that whatever the idiosyncracies of a management may be they can be given some kind of reflection in a matrix C which may have more or fewer rows than A.

Contrary to the matrix A in (4.1), this matrix C is completely arbitrary and can be modified, simplified, or complicated in any way that management wants. For example, management may require

only that total dollar sales be reported, in which case the matrix C consists of only the first row of the matrix in Example 1. Or they may want to know the total dollar sales of Product 1 and the physical sales volume of Product 2, in which case the matrix C consists of the second and third rows of the matrix in Example 1. In sum, then, our matrix C is to be prepared in such a way that it will supply the indicators that a management, for whatever reason, may want to obtain.

We now want to evaluate the effect of obtaining such indicators from the specific viewpoint of control, i.e. how these indicators may be used to supply more or less valid information on the value of goal variables v. We are not here concerned with other kinds of information requirements. Our only concern is as follows. Given a managerially prescribed set of goals v and a matrix C formed in accord with management's wishes, how can we proceed to secure information on the goals in v via the indicators w by reference to A, C and X?

We have referred to the components in w as indicators and we now emphasize two points concerning the characteristics of indicators. One is the fact that in order to know the performance in the goal attainment, it is not always necessary to measure the indicated variable directly. Surrogates 1/ may be used instead. For instance

1/ This term is adapted from Chapter XI in Charnes and Cooper, 1961. Note that even "profits" may be a surrogate for "general health of the enterprise."

the level of profit may be a goal but its levels need not be measured and reported directly if it may be inferred from the total dollar sales of the products. The latter may also be more immediately useful since, as we have mentioned, it has more "control measurability" than the profit level in the sense that profit level depends upon many other factors, including the total dollar sales, and it normally takes more time to calculate the profit level than to calculate the total dollar sales. This process is also capable of extension as witness, for instance, the widespread use of physical figures in place of dollars in reports for production-cost control. There is no reason, of course, why such physical reporting should be confined to production. Thus the total of dollar sales may be the goal of a sales manager without requiring the reports to be rendered only in these units, if, for example, total dollar sales can be inferred from the total shipment tonnage, the total number of orders, etc.

The second point we want to emphasize is the fact that the qualitative nature of indicators has next to nothing to do with their ability to contribute to a determination of the values of goal variables. As we remarked in Chapter I, we need not care what materials are used in a barometer but only whether its calibrations are sufficiently accurate to provide the weather indications that are wanted. Similar remarks apply to our indicators, too. That is, what we are concerned with in evaluating indicators is the relationship between what subgoal variables do on indicators and what they do on the given goal variables.

3. The Determination of Actual Values of Goal Variables

The remarks made at the close of our previous section were directed towards emphasizing that our analysis is directed only towards knowing the relations between subgoals and goals, as given by $Ax = v$, and how the indicators w in $Cx = w$ may be used to supply information on these relations. We now proceed in our analysis of these indicators by first raising the following questions. Under what conditions can we determine the value of the goal variable v by knowing the value of the indicators w?[1] That is, we assume that

[1] See, e.g., Wagner, 1962, for a treatment on a similar topic from a statistical sampling view point.

we know, beforehand, the matrices A and C, and that the value of w is supplied to us from time to time. But if we do not know x we cannot know v via $Ax = v$ and hence we cannot know the deviations in $v-b$.

The question under consideration now becomes: under what condition on A and C can we precisely determine v from w? In this section we shall assume that there are no constraints imposed upon subgoal variables, not even non-negativity conditions. Then, we shall introduce such constraints in the next section after simplifying matters in this one, so that we can more readily understand the relationship among goals, subgoals, and indicators.

Before going into a mathematical analysis the following simple example may be of help to clarifying it.

> Example 2: A firm produces two kinds of products, Product 1 and Product 2, whose contribution to overhead and profit are \$1 per unit and \$.5 per unit, respectively. The management wants to know the total contribution to overhead and profit from time to time.

Suppose that the gross weights of Product 1 and Product 2 are g_1 pounds per unit and g_2 pounds per unit, respectively and that the total weight of products shipped is reported as $g(=g_1x_1+g_2x_2$, where x_1 and x_2 are the sales quantities of Product 1 and Product 2, respectively.) to the firm's management from time to time--e.g., as part of a contract with an agent who handles all shipments.

Under what conditions on g_1 and g_2 can the management determine the actual amount of contribution to overhead and profit (or simply contribution level) from the total weight of products shipped?

Suppose that $g_1 = 2$ and $g_2 = 2$. Then, in this case, we cannot determine the contribution level from the knowledge of the weight of products shipped, say 1,000 pounds, since the contribution level may be anywhere between \$250 and \$500. Furthermore, this range for the contribution level was determined because we know the sales volumes of the two products never assumes negative values. If, however, the sales volumes had not been constrained to be non-negative, for any value of the total weight of products shipped then the data $g_1=2$, $g_2=2$ would supply even less information and, in fact, the contribution level could then be any real number.

Obviously, in this case, the necessary and sufficient condition for the determination of the contribution level from the total weight of product shipped is $g_1 = 2g_2$, since then the management can determine the contribution level they have achieved from the total weight of products shipped by dividing it by g_1. E.g., if the total weight of products shipped is 1,000 pounds where $g_1 = 4$ $g_2 = 2$, the management, perhaps via prior accounting manipulations, could determine the contribution level as exactly equal to \$250.

This means that the matrix A, (in this case (1, .5)), must be expressible as a scalar multiple of the row of C, in this case (g_1, g_2).

We shall extend our insight into the required relationship between A and C at the same time that we generalize upon it by means of the following theorem.

Theorem 1:

Let (4.1) $v = Ax$

and

(4.2) $w = Cx$

where A and C are $(m \times n)$ and $(k \times n)$ matrices, respectively, and $x \in E^n$. A necessary and sufficient condition for v to be a uniquely determined function of w only is that each row of A is expressible as a linear combination of the rows of C.

To prove this theorem we can proceed as follows. Let $C^{\dagger}$ be the $(n \times k)$ generalized inverse of C, and let C^{o} be an $n \times (n-r)$ matrix whose columns are basis vectors for the null space of C. Then, from (4.2) we obtain an expression for x,

$$x = C^{\dagger}w + C^{o}z \qquad (z \in E^{n-r}). \tag{4.3}$$

(Refer to Section 5 of Appendix A.) Next, from (4.1) and (4.3) we obtain

$$v = AC^{\dagger}w + AC^{o}z \qquad (z \in E^{n-r}). \tag{4.4}$$

In general v is not uniquely determined when w is specified because of the arbitrary vector z. On the other hand, a necessary and sufficient condition that v be uniquely determined as a function of w only can evidently be attained via the expression $AC^{o}z = 0$ for all $z \in E^{n-r}$, or simply $AC^{o} = 0$.[1]

---

[1]If $AC^{o}z$ is equal to a constant vector for all $z \in E^{n-r}$, v is also uniquely determined. But obviously such a constant vector has to be the zero vector since for $z = 0$, $AC^{o}z = 0$ for any A and C.

The above condition is equivalent to saying that $C^{o}z$ is a vector in the null space of A for all z in E^{n-r}. Note that $C^{o}z$ $(z \in E^{n-r})$ need not span the null space of A, but only that it must be included in the null space of A. However, $C^{o}z$, $(z \in E^{n-r})$ does span the null space of C, since, by definition, C^{o} is a matrix of a null space basis of C.

By virtue of the remarks just entered, we can now proceed

as follows. Let N(A) and N(C) represent, respectively, the null spaces of A and C. Then, the above condition for unique determination of v is equivalent to

(4.5) $$N(A) \supseteq N(C).$$

I.e., the null space of C is contained in the null space of A. The vectors of N(C) then form a subset of N(A) although not necessarily a proper subset. However, in general, an n-dimensional Euclidean space, E^n, can be written as a direct sum of two orthogonal and complementary subspaces as follows:

(4.6) $$E^n = N(A) \oplus R(A^*) = N(C) \oplus R(C^*),$$

where R(A*) and R(C*) represent the range of A* and C* which are the transpose of A and C, and $\oplus$ means "direct sum."[1] From (4.5)

[1]See Section 4 of Appendix A.

and (4.6), we obtain

(4.7) $$R(A^*) \subseteq R(C^*),$$

which is equivalent to saying that each column of A* (i.e., each row of A) must be expressible as a linear combination of the columns of C* (i.e., the rows of C). Q.E.D.

Remark 1: If C is non-singular — hence n x n — then the condition of Theorem 1 is automatically fulfilled since $C^o = 0$. This may also be seen from the fact that the rank of A is at most equal to n, the number of columns of A, while the rows of an n×n non-singular C span every point in the n-dimensional Euclidean space, E^n.

Let us call a set of indicators w whose relationship with the subgoal variables, given by the matrix C, satisfies the condition in theorem 1 a "perfect indicator set." While it is desirable to have such a "perfect indicator set"--since they uniquely specify the values of goal variables--it may be very expensive or even impossible to obtain. When this is the case, there is then some interest in seeing how a less than perfect indicator set may be employed and this is the task to which we will turn in the next section, not only by relaxing the conditions of theorem 1 but also by imposing constraints upon the subgoal variables.

4. Constraints on Subgoals

We now turn our attention to cases in which the subgoal variables x are constrained by conditions of the form

$$Bx \leq h, \tag{4.8}$$

where B is an s x n matrix and h is an s-component column vector.

To make our analysis more concrete, let us analyze the relationship between one goal variable v and one indicator w which are related by means of a set of n subgoal variables x as follows:

$$\begin{aligned} ax &= v \\ Bx &\leq h \\ cx &= w \end{aligned} \tag{4.9}$$

where a $(1 \times n)$, $B(s \times n)$, $c(1 \times n)$, and $h(s \times 1)$ are vectors or matrices, as indicated, while v and w are ordinary real numbers. We shall assume that a, B, and c, are all non-zero in order to exclude trivial cases and we shall also assume that (4.9) is solvable. Note that because of the constraint imposed upon x ($BX \leq h$) the goal variable v may also be limited to a certain range. Our problem is, however, how knowledge of the w indicators may be used to narrow the range of the goal variable v in its assessment values for b.

The range for v without any knowledge of the indicator w is given by the solutions to the following two linear programming problems:[1]

[1] Conditions such as $x \geq 0$, if present, are assumed to be included in the constraints $Bx \leq h$.

$$(4.10) \quad \begin{cases} \text{Maximize} & ax \\ \text{subject to} & Bx \leq h \end{cases} \quad \begin{cases} \text{Minimize} & ax \\ \text{subject to} & Bx \leq h \end{cases}$$

I.e., the solutions to these two problems will produce vectors $\hat{x}$ and $\check{x}$ which can be substituted in $v = ax$ in (4.9) to give upper and lower values, $\hat{v}$ and $\check{v}$, respectively, for the range of v. If the solution to one or both of the linear programming problems in (4.10) is infinite, the corresponding upper bound or lower bound or both does not exist for the range of v. In such cases, we may put an aribitrary upper or lower bound or both which is large enough so that it does not affect our analysis. This is done by means of what are called "regularization methods"[1] which can be used without affecting the analysis and losing generality in any linear programming problem.

[1] See Charnes and Cooper, 1961.

If, on the other hand, no solution exists for one of the problems in (4.10), a solution does not exist for the other problem in (4.10), and this implies that no subgoal variables x can satisfy all of the constraints (i.e., $Bx \leq h$) simultaneously. This, too, may be handled by regularization as well as by the kinds of incompatible multiple goal formulations that were discussed in the previous chapter. Our interest is now, however, in "control measure" problems and not the kinds of incompatibilities that were previously examined. Hence we shall simply exclude such unsolvable cases from consideration here.

It is now of interest to analyze some of the effects that may ensue from the use of an indicator. In general, this produces a narrower range as can be seen by substituting the expression for x in (4.3) into (4.10) to obtain the following two new problems:

$$(4.11) \qquad \begin{cases} \text{Maximimize} & ac^{o}z + ac^{\dagger}w \\ \text{subject to} & Bc^{o}z \leq h - Bc^{\dagger}w \end{cases} \qquad \begin{cases} \text{Minimize} & ac^{o}z + ac^{\dagger}w \\ \text{subject to} & Bc^{o}z \leq h - Bc^{\dagger}w \end{cases}$$

where $B(s \times n)$, and $c^{o}(n \times (n-1))$ are constant matrices; $a(1 \times n)$, $h\ (s \times 1)$, $c^{\dagger}\ (n \times 1)$ are constant vectors; and w is a scalar; while $z\ ((n-1) \times 1)$ is a variable vector. The range for v derived from (4.11) is at most equal to and ordinarily less than the range for v derived from (4.10), since the effect of substituting $x = c^{\dagger}w + c^{o}z$ into (4.10) is to cut the original convex set generated by a hyperplane given by $cx = w$ and to reduce the original convex set to points in the intersection of the hyperplane and the original convex set.[1]

[1] An alternative way of effecting such reductions is via the concepts of solution space, requirements space and representative convex sets as given in Charnes, Cooper, and Henderson, 1953.

Then each time we obtain a new indicator w_i the dimension of the convex set is reduced by one due to the cut by the hyperplane $c_i x = w_i$ provided that c_i is independent of the set of $c_1, c_2, \ldots, c_{i-1}$. Thus, the range for v becomes narrower and narrower each time we add a new independent indicator and finally v is uniquely determined

when we obtain a set of n independent indicators.

We may make the point clear by diagramming the situation for a very simple example as follows:

> Example 3: Consider the machine loading problem of Example 1 in Chapter III. One unit of Product 1, whose contribution to overhead and profit (or simply contribution) is $1, requires 3 hours of processing on "Machine 1" and 5 hours on "Machine 2"; whereas one unit of Product 2, whose contribution is $.5, requires 2 hours on "Machine 1" only for its processing.
>
> If the available hours are, in total, 12 hours for Machine 1 and 10 hours for Machine 2, the convex set generated by the machine hour limitations and the non-negativity condition of variables can be represented by the shaded area in the following Figure IV-1.

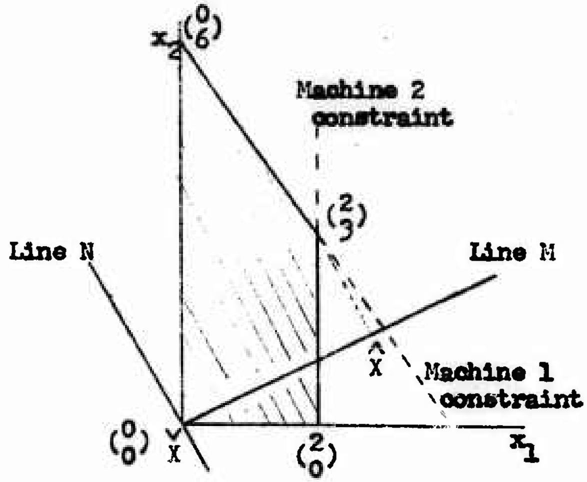


Figure IV-1

Feasible Machine Loading Program

The Line N here represents a set of points whose contribution to overhead and profit is zero,[1] i.e. Line N is the null space

[1]Its only intersection with the convex set of solutions occurs at $x_1 = x_2 = 0$.

of $a = (1, .5)$. (Refer to Section 4 of Appendix A.) Line M is orthogonal (or perpendicular) to Line N starting from the origin and represents the range of $a^* = \binom{1}{.5}$. (Again refer to the discussion on the null space of a matrix and the range of its transposed matrix in Section 4 in Appendix A.) The solution space[1] for $ax = v$

[1]This term is not used in exactly the same sense that has now become standard in the linear programming literature since the appearance of Charnes, Cooper, and Henderson, 1953.

for any given v is represented by a line parallel to Line N, since, as we have discussed in Section 3 of Chapter II,[1] if x^* is a solution

[1]Also refer to Section 5 of Appendix A.

to $ax = v$, then any vector $(x^* + x^o)$ is also a solution to $ax = v$ if and only if x^o is a vector in the null space of a. Therefore, the segment on Line M generated by the orthogonal (i.e., perpendicular) projection of the convex set onto Line M gives us the range for v. Let $\hat{x}$ be the point on this segment farthest from the origin and $\check{x}$ be the point on the segment closest to the origin as shown in Figure IV-1 above. Then the range for v is given by $a\hat{x}$ and $a\check{x}$,

i.e., $a\check{x} \leq v \leq a\hat{x}$ in this case.

Suppose that we have information on the amount of the total dollar sales from time to time. If p_1, the selling price of Product 1, is exactly twice p_2, the selling price of Product 2, — i.e., $p_1 = 2p_2$ — then management can easily determine the total amount of contribution by just dividing the total dollar sales by p_1. However, if $p_1 \neq 2p_2$, — i.e., if p_1 is any <u>other</u> constant multiple of p_2 — then this is no longer possible (except special cases at extreme points of the convex set, e.g. if the total dollar sales is zero we know the total contribution is also zero). From a practical standpoint, then, we now become interested in how an exact determination might be replaced by another "best" one in the sense of a "narrowest" range. In particular we now seek ways to narrow the range for the total contribution by the use of available information on the total dollar sales, even if we cannot uniquerely determine the total overhead-plus-profit contribution from a figure for total sales alone.

To illustrate, let us assume that $p_1 = p_2 = 2$. Then we have,

$$(4.12) \qquad w = cx = (2, 2)\, x$$

as an available indicator for the management. Of course, as mentioned earlier, we do not care about the qualitative nature of the indicator, whether it is the total dollar sales, or the total dollar consumption of packing materials, or the total weight products shipped. That is, we are to take whatever indicators are specified and translate them into information on v via $Ax = w$ and $Cx = w$.

From a reading of this indicator w in (4.12) we obtain $w = w_0$, which

implies that x must lie in the solution space of $cx = w_0$. This solution space is represented by Line S in the following diagram. However, because of the constraints imposed upon x we know that

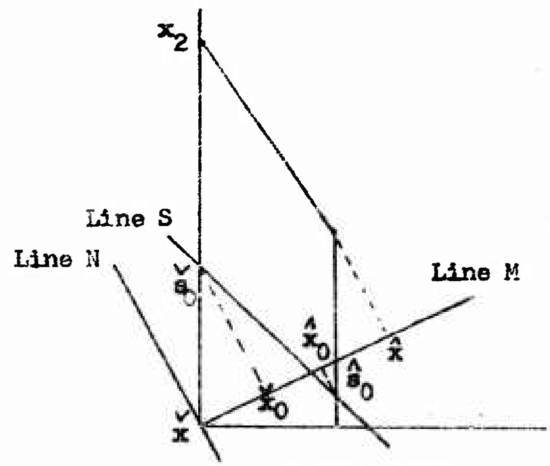


Figure IV-2
The Limitation on the Range of the Goal

x must lie on the solid line segement which connects $\check{s}_0$ and $\hat{s}_0$. Thus, the original convex set -- i.e., the shaded area in Figure IV-1 -- is now reduced to the line segment $\overline{\hat{s}_0 \cdot \check{s}_0}$ by knowledge of the total dollar sales.

Next, we let $\hat{x}_0$ and $\check{x}_0$ be points on Line M obtained by orthogonal projection of $\hat{s}_0$ and $\check{s}_0$, respectively, on M. Then the original range for v, which was between $\hat{x}$ and $\check{x}$, is now narrowed to $\hat{x}_0$ and $\check{x}_0$. Note, however, that if Line S is parallel to Line N -- as was to case when we had $p_1 = 2p_2$ -- then $\hat{x}_0 = \check{x}_0$ and we can determine v uniquely since $a\check{x}_0 = v = a\hat{x}_0$.

The discussion for this example suggests the desirability

of a chart depicting the relations between w and v that may be obtained by solving (4.11) for various values of w. Such a chart is given in Figure IV-3, where the upper boundary, which consists of three line segments, (1), (2), and (3), may be

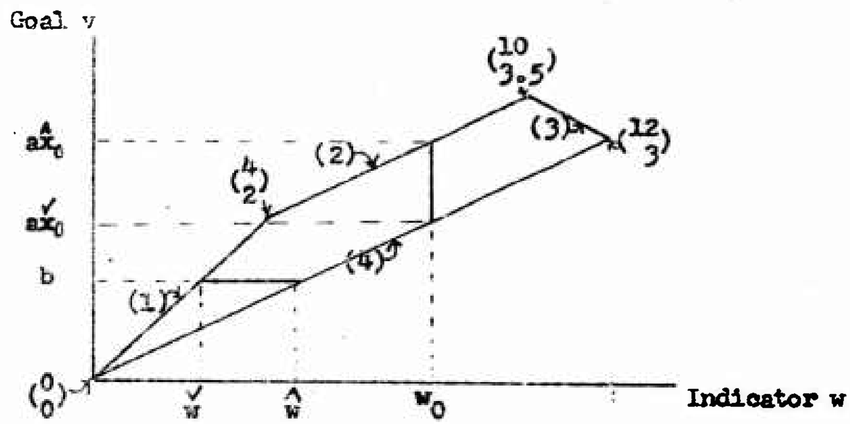


Figure IV-3

An Indicator-Goal Control Chart

called the "maximum contribution line," since it is the locus of $\hat{ax}_0$ for various values of w, i.e., the maximum obtainable contribution at any given total dollar sales, and the lower boundary, given by the segment (4), may be regarded as the "minimum contribution line" since it is the locus of $\check{ax}_0$ for various values of w, i.e., the minimum abtainable contribution at any given total dollar sales.

We may call this an "indicator-goal control chart" by virtue of the fact that if management observes w_0, they can be sure that the contribution level is at best $\hat{ax}_0$ and at worst $\check{ax}_0$ without

knowing the actual sales product mix. Furthermore, the goal parameter b (the target value of the contribution level) can also be interpreted (beforehand) in terms of the total sollar sales. In Figure IV-3, if the total dollar sales is less than or equal to $\check{w}$, we are sure that the goal level has not been attained whereas if the total dollar sales is greater than or equal to $\hat{w}$, we are sure that the goal level has been attained. In this sense, the segments (1) and (2) in the diagram may be called the "minimum dollar sales line", since it represents the minimum amount of dollar sales to achieve any given contribution level, whereas the segments (3) and (4) may be called the "maximum dollar sales line," since it represents the maximum amount of dollar sales to achieve any given contribution level.

Notice that the points, $\binom{0}{0}$, $\binom{4}{2}$, $\binom{10}{3.5}$, and $\binom{12}{3}$ in Figure IV-3 corresponds to the points, $\binom{0}{0}$, $\binom{2}{0}$, $\binom{2}{3}$, and $\binom{0}{6}$ in Figure IV-1. Actually the convex set in Figure IV-3 is a transformed or "mapped" convex set of the original one given Figure IV-1. Therefore, in order to prepare an indicator-goal chart like the one given in Figure IV-3, we need not find the maximum and the minimum value of v for all values of w, but only for some "critical" values of w which correspond to the extreme points in the "mapped" convex set.[1/]

1/ See Theorems 4 and 5 on p. 237 in Charnes and Cooper, 1961.

Note that, regardless of the dimension of the original convex set, the "mapped" convex set is always two dimensional. That is the latter is two dimensional because here we are interested in the original convex set only for its implications relative to

the relationship between the goal variable v and the indicator w. Of course a practical means of implementation is wanted if such indicator-goal control charts are to be used and so, in Appendix C we have presented an algorithm for generating all the extreme points of the "mapped" convex set of an original convex set of any dimension. (This algorithm has been derived by making suitable modifications of the simplex algorithm.)

A managerial application of such an indicator-goal control chart will be discussed in an experimental context in more detail in Section 5 of Chapter VI. Here, however, we may usefully pause and consider how Figure IV-3 may be used to cast light on possible conflicts between goals and subgoals which originate in information devices that might be used by management. Suppose, for instance, that managements wants to maximize v, the contribution to overhead and profit and to translate this into an operational subgoal. Then they tell one of their subordinates that his performance will be evaluated by the indicator w which is the total dollar sales. This is translated by subordinate, a sales manager, say, into an effort to maximize w. If this is true then, from Figure IV-3, the subordinate will operate at $\binom{12}{3}$ which brings the maximum w since it is the point furthest from the origin in the direction of w. But this is not the maximum v which occurs at $\binom{10}{35}$. We do not wish to go into an analysis of organizational factors involved in control, and hence we confine ourselves only to direct constraints that might be imposed on w. Clearly if we are to have an

effort at maximum w result in a choice of $\binom{10}{3.5}$ we must constrain w, in one way or another, so that all of the area of the convex set to the right of, and below, this point is eliminated. If this is done we can be sure that a maximum w and a maximum v will coincide with w = 10 and v = 3.5. This and related kinds of information, too, may be directly extracted from charts like Figure IV.3 which is not the case, of course, for ordinary breakeven and profit graph portrayals.

5 Choice of an Indicator

When we are going to choose an indicator, w, by which we want to measure the goal level, v, the choice may be made on an indicator which minimizes the area of the convex set in an indicator-goal control chart, or, alternatively, an indicator which minimizes the maximum range for v at particular values of w, etc. That is, the desirability of an indicator may be defined in various ways.

Note that if an indicator is a "perfect" indicator by which the vector a in the goal-subgoal relationship can be described as a scalar multiple of the vector c for the indicator, then the convex set is a straight line going through the origin. In general, however, such a choice is dependent upon the shape of the convex set, as well as the vectors a and c. Therefore, it may be inferred independently of the convex set involved in the problem so that an attempt for improving an indicator can be made independently of the convex set.

As a way to derive such a criterion, consider the following quantity

(4.13) $$\gamma = \sqrt{1 - \left(\frac{ac^*}{|a| \cdot |c|}\right)^2} \qquad (a \neq 0,\ c \neq 0)$$

where $|a| = \sqrt{aa^*}$, $|c| = \sqrt{cc^*}$.

In our example, since $a = (1,\ 5)$ and $c = (2,\ 2)$, γ is obtained from

(4.14) $$\gamma = \sqrt{1 - \left(\frac{1 \times 2 + .5 \times 2}{\sqrt{1^2 + .5^2}\sqrt{2^2 + 2^2}}\right)^2} = \sqrt{\frac{10 - 9}{10}} = .316$$

Note that if a is a scalar multiple of c, i.e., the indicator is "perfect," then $\gamma = 0$. Actually γ is an angualr measure.[1] It is,

[1] This is also true of other commonly employed measures such as the so-called Pearsonian correlation coefficient in statistics.

in fact, sine of the angle between the range of a* and the range of c*, since the quantity $\frac{ac^*}{|a|\cdot|c|}$ is the cosine of the angle between the range of a* and the range of c*. Therefore, if c* is included in the null space of a, we will have a c* = 0 and hence $\gamma = 1$. We shall call γ an "indicator-goal divergence coefficient" or simply a "divergence coefficient."[1] This

[1] Alternatively, we may call $\hat{\gamma} = \sqrt{1 - \gamma^2} = \frac{ac^*}{|a|\cdot|c|}$ an "indicator-goal conformity coefficient" or simply a "conformity coefficient."

divergence coefficient γ may be interpreted as follows: Consider a convex set given by

(4.15) $$x^*x = x_1^2 + x_2^2 + \ldots, + x_n^2 \leq 1$$

Then γ is half of the maximum range for v due to the fact that R(c*) and N(c) are orthogonal as shown in the following diagram. Note that the range for v is maximum when the solution set for cx = w subject to $x^*x \leq 1$ is included in N(c), i.e., when w = 0.

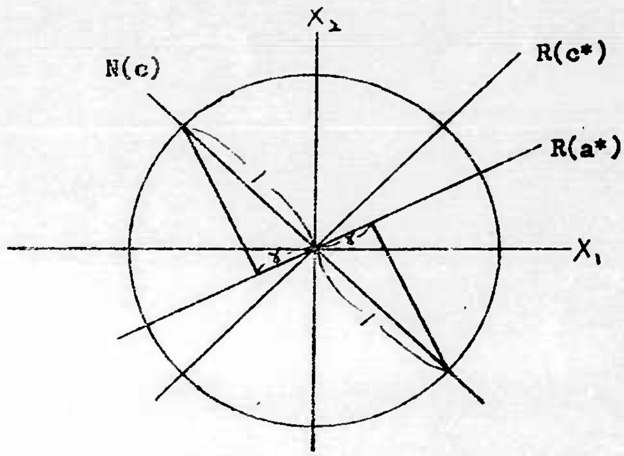


Figure IV-4

Indicator-Goal Divergence Coefficient

The divergence coefficient is intended to characterize the relationship between the subgoals and the indicator (represented by the vector c) in terms of the relationship between the subgoals and the goal (represented by the vector a). The actual width of v is determined by the shape of the convex set involved in the problem and the actual value of the indicator reading as well as the divergence coefficient.

As we have seen, γ corresponds to a cosine value for the angle between the ranges of a* and c* with $\gamma = 0$ meaning that an indicator is perfect. But it does not follow that $\gamma = 1$ implies that the indicator does not help in narrowing the range for v. It only means that $\gamma = 1$ is at its worst possible value. For example, in Figure IV-5, below, which is reproduced from Figure IV-2, if we take c = (1, -2), then R(c*), the range of c*, is completely included in Line N which is N(a) the null space of a. Similarly Line M which is the range of a is completely included in N(c), the null space of c. Here N(c) is

$(\frac{1}{5})a$ where a is a scalar. As discussed in Appendix A, a solution space to $cx = w$ for any given value of w is given by a parallel shift of $N(c)$. But then the intersection of the convex set and a line parallel to Line M ($= N(c)$) produces a line segment whose points satisfy the constraints imposed upon x and $cx = w$ at a given value of w. Thus in this case let w_0 be such a w. By projecting the line segment perpendicularly onto Line M, it will be noted that the original range $(\check{x}, \hat{x})$ is narrowed to $(\check{x}_0, \hat{x}_0)$. Hence, in this case for any value of w we obtain a narrower range even for an indicator whose divergence coefficient takes even the worst possible value, i.e., 1.

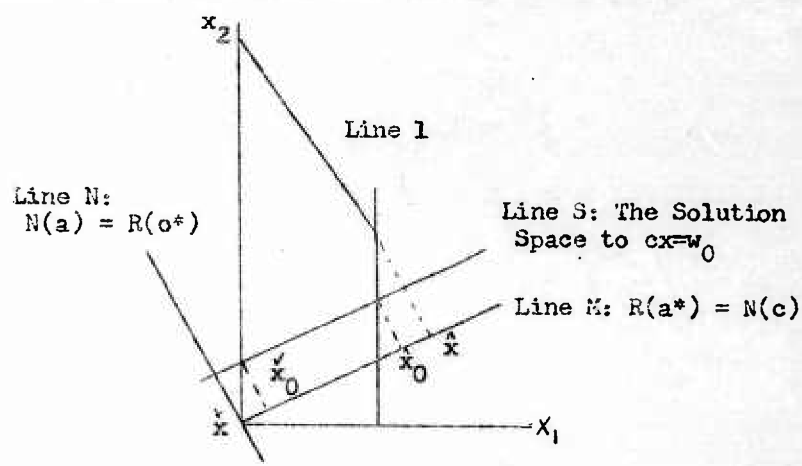


Figure IV-5

An Indicator with $\gamma = 1$

Actually, when constraints are imposed upon x <u>any</u> indicator produced by $cx = w$ helps narrow the range of v for at least <u>some</u> values of w, unless (i) the indicator can be produced by

a linear combination of some other indicators whose values are already known[1] or (ii) the convex set is given by <u>only</u> two hyperplanes both parallel to H(a), the convex set being unbounded.

[1] Note that the convex set given in Figure IV-5 may be considered as a subset of the original convex set produced by a "cut" by some other indicators.

Therefore, a precise characterization of an indicator with respect to its goal can only be made taking into account the set of constraints imposed upon the subgoal variables. However, it may still be desirable to have a factor by which an indicator is characterized in terms of the goal only, since the choice of an indicator may have to be made without knowing the precise shape of the convex set. Pursuit of this topic, however, would require us first to undertake a discussion of conditions under which proper characterizations could be made of such unknown--i.e., imprecisely known--convex sets and to avoid such a further treatment, which would have to be rather extended, we close here now only with the following conjecture. Given any bounded convex set, an indicator with a smaller divergence coefficient will produce a mapped convex set in an indicator-goal control chart with a smaller area. This is meant to hold even when the original set is not known precisely, of course, for any such restrictions will then enable us to confine our attention to a smaller space of study for control purposes.

CHAPTER V

A SPREAD SHEET APPROACH TO ANALYSIS OF GOALS

1. Introduction

In this chapter, we want to explore the possibilities of applying the goal analyses that were developed in the previous three chapters. Specifically we want to do this in a way that will immediately connect our approach to the data that are ordinarily available under existing double-entry bookkeeping system routines. Our purpose in doing this is twofold. First, we want to preserve the rather considerable advances that have already benefitted management by means of the double-entry principle. Second, we want to do this in a way that also opens a prospect for continual progress by connecting the mechanics and the ideas of double entry analysis and record keeping to major parts of mathematical theories and methods.[1]

[1] See, e.g., the eloquent appeal for more accounting research to be directed along these lines in Mautz, 1963. See also Charnes, Cooper, and Ijiri, 1962.

First of all, we shall review the basic mechanism of the double-entry bookkeeping in Section 2. Then in Section 3, we shall explore the relationship between transactions and asset-equity balances by means of (i) a spread sheet, (ii) an incidence matrix, and (iii) an accounting network associated with an incidence matrix. In Section 4, as a way of combining the double-entry

bookkeeping system with the linear programming technique, we shall develop a spread sheet approach to goal analysis by using an example. Finally, in Section 5 we shall discuss the problems of aggregation of business activities in connection with the characteristics of "spread sheet planning."

2. The Double-Entry Bookkeeping System

The heart of the double entry bookkeeping system, which the nineteen century mathematician, Arthur Cayley, called one of the two perfect sciences[1], is the

[1] "The Principles of Book-keeping by Double Entry constitute a theory which is mathematically by no means uninteresting: it is in fact like Euclid's theory of ratios an absolutely perfect one, and it is only its extreme simplicity which prevents it from being as interesting as it would otherwise be." Cayley, 1907, preface.

idea that the total value of assets which a firm holds may be categorized or partitioned from two different view points; (i) by the type of the assets, such as cash, inventories, machinery, etc., and (ii) by the claimants on the total value of the assets, such as trade creditors, banks, stockholders, etc. Each partitioned part of the total value of assets is labeled as an "account." Accounts used in the former partition (or the "asset partition") are called "asset accounts" and in the latter partition (or the "equity partition"), "equity accounts."

Let A be an asset account and E be an equity account. Also let $\mathcal{A}$ be the set of all asset accounts used in an asset partition and $\mathcal{E}$ be the set of all equity accounts used in an equity partition. I.e.,

(5.1) $$\mathcal{A} = \{ A \}$$

(5.2) $$\mathcal{E} = \{ E \}$$

Let Y_A be the amount assigned to an asset account and Y_E be the amount assigned to an equity account. Also let $Y_{\mathcal{A}}$ be the

sum of all Y_A and $Y_{\mathcal{E}}$ be the sum of all Y_E. Then, since the asset partition and the equity partition are only two different ways of partitioning the same total value of assets, we have by definition,

$$(5.3) \qquad \sum_{A \in \mathcal{A}} Y_A = Y_{\mathcal{A}} = Y_{\mathcal{E}} = \sum_{E \in \mathcal{E}} Y_E.$$

This relation is, in fact, an identity or can be made to be one by suitable rules of valuation and summation. For instance, as in accounting in the United States, and elsewhere, the total value of a firm's assets is resolved into a total of dollar claims further partitioned between outside creditor claims and the claims of owners of the firm, both representing equity accounts. To make (5.3) an identity the owners' claim to the total value of the assets is determined as a residual of the total value of the assets, less the claim by the outside creditors. Furthermore, there is no non-negativity imposed, per se, on the owners' claim -- or even some of the other equities — and hence the resulting "equality" is always brought about. That is, under these rules, the expression (5.3) becomes an identity since it always holds.

Now we next introduce the idea of a transaction which we designate as any action that can change any one (or more) of Y_A or Y_E ($A \in \mathcal{A}$; $E \in \mathcal{E}$). From this standpoint the action can be anything, even a windstorm, a fire, etc., if it can produce any such change. Thus, let Y_A^+ and Y_A^- be the amount of increase and decrease, respectively, of an asset account A ($A \in \mathcal{A}$) caused by a transaction and let Y_E^+ and Y_E^- be the amount of increase and decrease, respectively, of an equity

account E ($E \in \mathcal{E}$) caused by a transaction. Then, for any transaction, we have, from (5.3) and these definitions

$$\sum_{A \in \mathcal{A}} (Y_A^+ - Y_A^-) = \sum_{E \in \mathcal{E}} (Y_E^+ - Y_E^-) \tag{5.4}$$

which upon rearranging terms gives:

$$\sum_{A \in \mathcal{A}} Y_A^+ + \sum_{E \in \mathcal{E}} Y_E^- = \sum_{A \in \mathcal{A}} Y_A^- + \sum_{E \in \mathcal{E}} Y_E^+. \tag{5.5}$$

Elements on the left hand side of (5.5) are called "debit entries" and elements on the right hand side of (5.5) are called "credit entries." Thus, (5.5) may be interpreted to mean that any collection of transaction must produce a sum in which the debit entries and the credit entries are equal. The equation (5.3) may be called the fundamental equation of the double entry bookkeeping system from which (5.5) is derived by imposing further suitable definitions.

A transaction may give rise to more than one debit or credit entry. Such a transaction may be called a "compound transaction," in order to distinguish it from a "simple transaction" which gives rise to only one debit and credit entry. However, any compound

transaction may be decomposed into a set of simple transactions by combining each debit entry with each credit entry in a proportionate amount. Suppose, for instance, that a compound transaction gives rise to m debit entries whose amounts are represented by d_1, d_2, ..., d_m, respectively, and n credit entries whose amounts are represented by c_1, c_2, ..., c_n, respectively. Then, the compound transaction may be decomposed into a set of m x n simple transactions, each one of which relates the i^{th} debit entry ($i = 1, 2, \ldots, m$) with the j^{th} credit entry ($j = 1, 2, \ldots, n$) by the relations $d_i \dfrac{c_j}{\sum_{k=1}^{n} c_k} = c_j \dfrac{d_i}{\sum_{k=1}^{m} d_k}$. Hence, in the following discussions, we can, without loss of generality, deal with only simple transactions, on the assumption that such decompositions are available and applied when compound transactions are involved.

Note that the above fundamental equation of the double entry bookkeeping does not require a single unit of measurement in aggregating assets and equities, since Y_A and Y_E in (5.3) may be interpreted as vectors which represent a set of characteristics of assets and equities measured in individual units, e.g., in dollars, in bushels, in production capacities, etc.

We shall develop this point in somewhat more detail as follows. Let $\vec{Y}_A$ be a row vector for a set of characteristics of an asset account and let $\vec{Y}_E$ be a row vector for a set of characteristics of an equity account. Also, let $\vec{Y}_{\mathcal{A}}$ and $\vec{Y}_{\mathcal{E}}$ be row vectors of a set of characteristics of aggregated assets and equities, respectively, i.e., $\vec{Y}_{\mathcal{A}} = \sum_{A \in \mathcal{A}} \vec{Y}_A$ and $\vec{Y}_{\mathcal{E}} = \sum_{E \in \mathcal{E}} \vec{Y}_E$. In addition, let $[Y_{\mathcal{A}}]$ and $[Y_{\mathcal{E}}]$ be matrices of a set of all $\vec{Y}_A$ $(A \in \mathcal{A})$ and a set of all $\vec{Y}_E$ $(E \in \mathcal{E})$, respectively. I.e., each row of $[Y_{\mathcal{A}}]$ represents a $\vec{Y}_A$ and each row of $[Y_{\mathcal{E}}]$ a $\vec{Y}_E$. Suppose that we have m assets accounts and n equity accounts, all of which are represented by their ℓ characteristics measured in respective units. Then, $[Y_{\mathcal{A}}]$ is an $m \times \ell$ matrix and $[Y_{\mathcal{E}}]$ is an $n \times \ell$ matrix. We now state the "multi-demensional" fundamental equation of the double entry bookkeeping system as follows.

$$(5.3a) \qquad \sum_{A \in \mathcal{A}} \vec{Y}_A = e^*_m [Y_{\mathcal{A}}] = \vec{Y}_{\mathcal{A}} = \vec{Y}_{\mathcal{E}} = e^*_n [Y_{\mathcal{E}}] = \sum_{E \in \mathcal{E}} Y_E$$

where e^*_m and e^*_n are m-component and n-component, respectively, row vectors whose components are all unity.

Similarly, the equation (5.5) which was derived from (5.3) may be extended to

$$(5.5a) \qquad \sum_{A \in \mathcal{A}} \vec{Y}_A^{+} + \sum_{E \in \mathcal{E}} \vec{Y}_E^{-} = \sum_{A \in \mathcal{A}} \vec{Y}_A^{-} + \sum_{E \in \mathcal{E}} \vec{Y}_E^{+},$$

where $\vec{Y}_A^{+}$ and $\vec{Y}_A^{-}$ are vectors of the amounts of increase and decrease, respectively, of characteristics of an asset account caused by a transaction and $\vec{Y}_E^{+}$ and $\vec{Y}_E^{-}$ are vectors of the amounts of increase and decrease, respectively, of characteristics of an equity account caused by the transaction.

Here, again, the fundamental equation (5.3a) may be viewed as an identity by defining the vector for the owners' equity, denoted by $\vec{Y}_C$, as a vector of residuals of each characteristic of aggregated assets, less each characteristics of aggregated claims by outside creditors, and allowing components of the vector for the owners' equity to be negative if necessary.

For example, suppose a business man invested $10,000 in cash for his business, borrowed $5,000 from a bank, and then purchased 4,000 bushels of wheat at $3 a bushel. Under the current accounting system, a balance sheet after these transactions will be as follows:

Table V-0

Ordinary Balance Sheet

Cash	$ 3,000	Loans	$ 5,000
Wheat	12,000	Proprietorship	10,000
	$15,000		$ 15,000

However, it is perfectly legitimate, under the double entry principle, to write a balance sheet in multiple units as follows:

Table V-0a

Balance Sheet in Multiple Units

Cash	(3,000, 0)	Loans	(5,000, 0)
Wheat	(0, 4,000)	Proprietorship	
		Capital stock	(10,000, 0)
		Retained Earnings	(-12,000, 4,000)
		Total Proprietorship	(-2,000, 4,000)
	(3,000, 4,000)		(3,000, 4,000)

Here, first elements in parentheis are stated in dollars and the second elements in bushels of wheat. Since, the bank loan must be paid in cash, $5,000 is entered as a first component of Loans account. (If the loan must be paid by wheat, the loan is recorded as a second component of Loans account.)

Observe, then, the following significance of the balances of the proprietorship accounts. First, the balances of the <u>total</u> proprietorship show how much the proprietor owns in net (if the component is positive) and how much be owes (in net) outside creditors (if the component is negative), after canceling all assets with claims by outside creditors, component by component. In the above example, the proprietor owns 4,000 bushels of wheat but has a shortage in cash by $2,000 if the loan is required to be paid immediately, i.e., he must manage the shortage in cash by converting a part of 4,000 bushels of wheat into cash.

Therefore, the balances of the total proprietorship account, i.e., $\vec{Y}_C$ defined above, summarized the net balances of assets and equities, component by component, for their characteristics which we want to measure.

Consider, next, the balances of retained earnings account, i.e., $\Delta \vec{Y}_C$. Each component of $\Delta \vec{Y}_C$ represent a net increase or a net decrease of a characteristic of assets and equities resulted from the operations in a period. In the above example, it shows the fact that wheat is increased by 4,000 bushels and cash is decreased by \$12,000.

If we want to determine the net profit for a period, we can do so by multiplying $\Delta \vec{Y}_C$ by a ℓ-component column vector p which is a vector for evaluation and aggregation of characteristics of assets and equities. For example, if $p = \binom{1}{3}$, $\Delta \vec{Y}_C\, p = (-12{,}000,\ 4{,}000)\binom{1}{3} = 0$, which means that there is no profit for the period. But, if $p = \binom{1}{4}$, $\Delta \vec{Y}_C\, p = 4{,}000$ showing the net profit of \$4,000 for the period. In the same way, the net value of assets may be given by $\vec{Y}_C\, p$. E.g., if $p = \binom{1}{3}$, $\vec{Y}_C\, p = (-2{,}000,\ 4{,}000)\binom{1}{3} = 10{,}000$ which is equal to the amount initial investment since no profit has been attained if $p = \binom{1}{3}$.

From the above discussion it can be seen that the double entry principle is preserved, without modification, even when we drop the usual assumption of dollar homogenization and its related fungibility properties. Moreover, the generalization preserves all aspects of the current practice of uniform dollar reductions as in financial reporting in the U. S. and elsewhere. This can be seen if we develop the latter as a special kind of reduction to bilinear forms as follows. First we alter the fundamental equation (5.3a) to

$$(5.3b) \qquad e^*_m [Y_{\mathcal{A}}] p = e^*_n [Y_{\mathcal{E}}] p,$$

where p is a price vector and e^*_m and e^*_n are "totalizers" so that the result is one in which a scalar occurs on each side of (5.3b) with total assets (in dollars) equal to total claims (also in dollars).

Of course, (5.3b) is not the only way of effecting such a reduction. Consider, for instance, vectors q^*_m and q^*_n, respectively, which are m-component and n-component row vectors, respectively, with differing weights for their elements and write:

$$q^*_m [Y_{\mathcal{A}}] \text{ and } q^*_n [Y_{\mathcal{E}}].$$

We need no longer require equality between the components of $q^*_m [Y_{\mathcal{A}}]$ and $q^*_n [Y_{\mathcal{E}}]$, respectively, in cases where we wish to consider deviations from the standpoint of their weighted significance, element by element, in the resulting vectors. On the other hand we can determine positive and negative adjustments by suitable rules which will bring such equality about. Alternatively we can prescribe suitable value weights, p, and obtain the indicated equality via

$$(5.3c) \qquad q^*_m [Y_{\mathcal{A}}] p = q^*_n [Y_{\mathcal{E}}] p$$

in which event p will assume the role of conversion values, as needed,

to bring about a balance in the accounts.

In summary, we have shown that the double entry principle holds even if we do not have a single measuring unit, such as a monetary unit, in order to emphasize thereby some of the generality which the double entry principle admits via suitable mathematical extensions.

In Chapter II-IV, we have analyzed the relationships among goals, subgoals, and indicators without restricting our attention to goals stated in monetary unit. These, too, as we have just shown, can be given reflection in double entry treatments even in cases where we want, perhaps, to replace monetary measures but more fundamental units such as the "satisfaction measures," that might be associated with goal attainments and discrepancies. Pursuit of the latter topic would, however, require prior consideration of other topics like utility measurement for individuals or organizations and the related mathematical properties of partial orderings which matrices (like vectors) admit. This would be too great a diversion from our main stream of analysis, however. Hence, remaining part of the thesis is directed, instead, toward analysis of the current--i.e., dollar value--double entry bookkeeping system with special reference to more effective uses of data it generates.

3. A Spread Sheet, an Incidence Matrix, and a Network Representations of the Double-Entry Bookkeeping System

Dealing only with simple transactions we now explore relationships between the balances of asset and equity accounts as are brought about by transactions and direct our discussions to (i) a spread sheet approach, (ii) an incidence matrix approach, and (iii) an accounting network approach.

1. A Spread Sheet Approach

Let $\mathcal{S}$ be an ordered array of asset and equity accounts in which m asset accounts are positioned first and then n equity accounts are positioned next, i.e.,

$$\mathcal{S} = \{\mathcal{A} ; \mathcal{E}\} = \{A_1, A_2, \ldots, A_m; E_1, E_2, \ldots, E_n\} \tag{5.6}$$

Let s be the number of asset and equity accounts, i.e., $s = m + n$. Then, by the "k^{th} account" $(k = 1, 2, \ldots, s)$ we shall mean the k^{th} account in $\mathcal{S}$, i.e., A_k if $k \leq m$ or E_{k-m} if $k > m$.

We now represent by $\check{u}_k$ and $\hat{u}_k$ the beginning balance and the ending balance, respectively, of the k^{th} account for $k \leq m$, and by $-\check{u}_k$ and $-\hat{u}_k$ the beginning and the ending balance, respectively of the k^{th} account for $k > m$. That is, $\check{u}_k$ and $\hat{u}_k$ are <u>non-negative</u> for asset accounts and <u>non-positive</u> for equity accounts.

For simplicity, we use matrix and vector notation as follows: Let $\check{u}$ and $\hat{u}$ be s-component column vectors whose elements are $\check{u}_k$ and $\hat{u}_k$, respectively, $(k = 1, 2, \ldots, s)$, and let

(5.7) $\Delta u = \hat{u} - \check{u}$.

Then, the basic equation (5.3) of the double entry bookkeeping system may also be represented by,

(5.8) $e^*\check{u} = e^*\hat{u} = e^*\Delta u = 0$,

where e^* is the transpose of e which is an s-component column vector with unity for all of its entries so that

(5.8a) $e^*\Delta u \equiv \sum_{k=1}^{s} \Delta u_k = 0$

and the accounts are in balance -- i.e., (i) $e^*\check{u} = 0$ means that the beginning balances are in order and (ii) $e^*\hat{u} = 0$ means that the ending balances are in order since $\hat{u}_k \geq 0$ for all $k \leq m$ while $\hat{u}_k \leq 0$ for $k > m$ and hence $e^*\hat{u} = 0$ means $\sum_{k=1}^{m} \hat{u}_k = -\sum_{k=m+1}^{s} \hat{u}_k$. Finally, (iii) $e^*\Delta u = 0$ means $\sum_{k=1}^{s} \hat{u}_k = \sum_{k=1}^{s} \check{u}_k$ which is necessarily true if $\sum_{k=1}^{s} \hat{u}_k = 0$ and $\sum_{k=1}^{s} \check{u}_k = 0$. Then, even though (5.8) maintains an equivalence with (5.3) it adds further information on the relations that must maintain to obtain the indicated equivalence at two different points in time.

We next represent by w_{ij} the total amount of simple transactions whose debit entries are all to be made to the i^{th} account and whose credit entries are all to be made to the j^{th} account. Then W, which is an s x s square matrix whose elements are w_{ij}'s, becomes the so-called "spread sheet" of double-entry accounting.[1]

[1]Refer to Kohler, 1952, as well as Charnes, Cooper, and Ijiri, 1962. Also, refer to the "balanced-margin tables" by Rosenblatt, 1956, or Kemeny, Schleifer, Snell, and Thompson, 1962, etc., for other matrix approaches to accounting. An earlier attempt by Cronhelm is discussed in Jackson, 1956, pp. 310-12.

We now want to relate W to the fundamental relation (5.3) which we do by relating W and Δu as in

(5.9) $$(W - W^*) e = \Delta u$$

where W^* is the transpose of W. This is permissible since W is square and hence $(W - W^*)$ is conformable. Also the indicated relation to the Δu of (5.7) obtains because an element in the i^{th} row and the j^{th} column of W represents the amount to be debited to the i^{th} account and credited to the j^{th} account. Therefore, $W_i e$, the sum of the elements in the i^{th} row of W, is simply the total of all debit entries to the i^{th} account. Similarly, the sum of the elements in the i^{th} column of W (given by the i^{th} element in W^*e) shows the total of all credit entries to the i^{th} account. Therefore, the difference $(We - W^*e)$ shows the net debit entry (if the component in $\Delta u > 0$) or the net credit entry (if the component in $\Delta u < 0$) to each one of accounts.

The matrix $(W - W^*)$ is skew-symmetric with diagonal elements all equal to zero.[1]

[1]This is related to the idea of hollow transformation matrices in so-called Leontief analysis. See Leontief, 1951.

Example 1: Consider the following simplified balance sheet of a firm at the beginning of a period.

Table V-1

Beginning Balance Sheet

1.	Cash	\$30	4.	Equity	\$50
2.	Finished Goods	10		(Stockholders'	
3.	Materials	10		account only)	
		\$50			\$50

If we arrange these account balances in the order of Cash, Finished Goods, Materials, and Equity, we have,

$$\check{u} = \begin{bmatrix} \check{u}_1 \\ \check{u}_2 \\ \check{u}_3 \\ \check{u}_4 \end{bmatrix} = \begin{bmatrix} 30 \\ 10 \\ 10 \\ -50 \end{bmatrix}$$

Suppose the firm had the following transactions during the period.

Purchases of Materials in cash	$w_{31} = 5$
Consumption of materials	$w_{23} = 2$
Fixed operating expenses paid in cash	$w_{41} = 2.5$
Cost of cash sales of finished goods	$w_{12} = 3$
Gross profit on cash sales of finished goods	$w_{14} = 3$

Here $s = 1, 2, 3, 4$ and so W is 4 x 4 and for these transactions, the "spread sheet" is

$$W = \begin{bmatrix} 0 & 3 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 5 & 0 & 0 & 0 \\ 2.5 & 0 & 0 & 0 \end{bmatrix}$$

Transposing and forming the expression (5.9), we have

$$(5.10) \qquad (W - W^*)e = \begin{bmatrix} 0 & 3 & -5 & .5 \\ -3 & 0 & 2 & 0 \\ 5 & -2 & 0 & 0 \\ -.5 & 0 & 0 & 0 \end{bmatrix} e = \begin{bmatrix} -1.5 \\ -1 \\ 3 \\ -.5 \end{bmatrix} = \Delta u,$$

from which we see the effects produced by these transactions in going from $\check{u}$ to $\hat{u}$. Alternatively using the data of $\check{u}$ as already given we apply (5.7) to obtain $\hat{u}$ by means of

$$(5.11)\quad \hat{u} = \check{u} + \Delta u = \begin{bmatrix} 30 \\ 10 \\ 10 \\ -50 \end{bmatrix} + \begin{bmatrix} -1.5 \\ -1 \\ 3 \\ -.5 \end{bmatrix} = \begin{bmatrix} 28.5 \\ 9 \\ 13 \\ -50.5 \end{bmatrix}.$$

Note that $(W - W^*)$ is skew symmetric and $\check{u}$, $\hat{u}$ and Δu all satisfy (5.8).

2. An Incidence Matrix Approach

The relationship between transactions and account balances may be analyzed by means of an incidence matrix.[1] This may also

[1] Refer to Charnes and Cooper, 1961, for more detail discussion on an incidence matrix.

be related to ideas of graphs and networks in mathematics as further points of contact between mathematics and accounting if we proceed as follows.

Let w be an $s(s-1)$-component column vector whose elements are w_{ij}'s ($i \neq j$; $i = 1, 2, \ldots, s$; $j = 1, 2, \ldots, s$). We shall use a "net method" of development in that we will eliminate all entries whose debit entry account and credit entry account are the same, i.e., the diagonal elements in the matrix W. Let T be an $s \times s(s-1)$ "incidence matrix" -- i.e., a matrix in which each column has only a 1 and -1 in two of its rows and zeros in all other entries. If the k^{th} element of the vector w is w_{ij}, then the k^{th} column of T has the entry 1 in the i^{th} row and the entry -1 in the j^{th} row with all other elements equal to zero.

Thus, by means of these definitions we can relate T, w, and Δu by

$$\text{(5.12)} \qquad Tw = \Delta u.$$

Of course the number of columns of T and the rows of w may be made smaller than s(s-1) when the matrix W (a spread sheet) has zero entries in off-diagonal cells.

Example 2: By using the data in Example 1 above, we may prepare a vector w and a matrix T as follows:

$$w = \begin{bmatrix} w_{31} \\ w_{23} \\ w_{41} \\ w_{12} \\ w_{14} \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 2.5 \\ 3 \\ 3 \end{bmatrix}$$

$$T = \begin{bmatrix} -1 & 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}$$

Then, we have

$$\text{(5.13)} \qquad Tw = \Delta u = \begin{bmatrix} -1.5 \\ -1 \\ 3 \\ -.5 \end{bmatrix}$$

which agrees whith the result in Example 1. Note that the non-zero entries in each column of T are given by the indices for each w_{ij}. E.g., w_{31} means that the first column of T has the entry 1 in row 3 since the first subscript of w_{31} is 3 and the entry -1 in row 1 since the second subscript of w_{31} is 1.

3. The Network Implication of the Incidence Matrix Approach

We now extend the above incidence matrix approach to an accounting network[1] in the following way in order to establish,

[1]Refer to Charnes, Cooper, and Ijiri, 1962. Also refer to Charnes and Cooper (1961) for more detail discussions on the network theory.

rigorously, how a graph-theoretic or network representation may be secured for the basic relationship between transactions and asset-equity balances in the double-entry bookkeeping system.

Any incidence matrix can be uniquely associated with a network. Thus, if T is an incidence matrix -- i.e., a matrix for which each column has a 1 and a -1 in two of its rows and zero entries everywhere else in the column -- we obtain a network as follows. A node is assigned to each row of T and these are then connected by branches in accordance with the positions (or incidences) of the non-zero entries, column by column, on these nodes. If the k^{th} column of T has the entry 1 in the i^{th} row and the entry -1 in the j^{th} row, the k^{th} branch is incident on the i^{th} node and the j^{th} node. Furthermore, we can orient these connections by means of arrows which we here do by pointing the arrowhead towards the positive incidence number, 1, and hence away from the incidence number, -1, for each node pair.

To produce a correspondence with our asset-equity balance and transaction ideas, we now interpret the nodes as asset or equity accounts and assign the branches to record all flows between them.

Then by choice of a suitable mathematical principle we also obtain access to the principles of double entry accounting.[1]

[1]See Charnes and Cooper, 1961 for precise details, including a discussion of the extremal conditions which are needed to obtain precise network characterization.

Example 3: The accounting network which corresponds to the incidence matrix T in Example 2 is as follows:

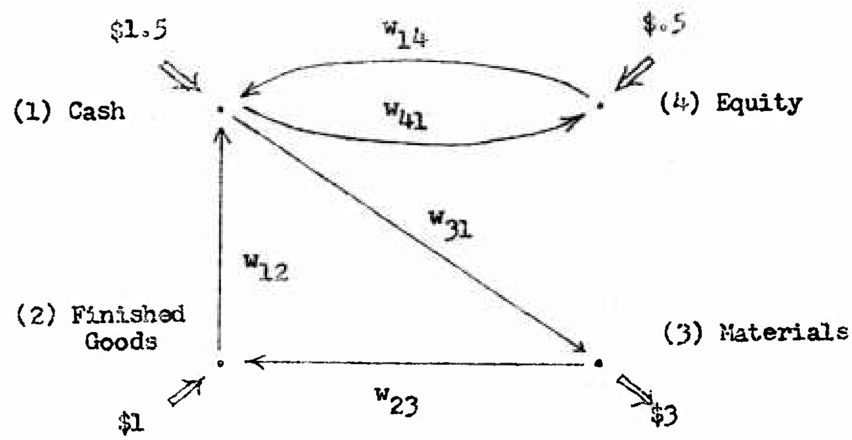


Figure V-1

An Accounting Network I

The short arrowheads at each node represent increase or decrease in the account balances. If we view these quantities as the conductance at each nodes, the above accounting network, and consequently the equation (5.12), may be considered as representing the Kirchhoff-node condition for the conservation of the current, or, equivalently, the preservation of the principle of double entry accounting.

We now want to utilize the above relationship between transactions and asset-equity balances represented through a spread sheet, an incidence matrix, and an accounting network as a means to obtain more managerially useful accounting information by such mathematical techniques as linear programming, generalized inverses, etc., which have been described in the previous chapters. Section 4 of this chapter and much of the next chapter are devoted to this purpose.

4. Analysis of Goals via Spread Sheet: An Example

In the previous section, we have shown the relationship between transactions and asset-equity balances by means of a spread sheet, an incidence matrix, and an accounting network. We now focus our attention on elements in a matrix W (a spread sheet) or, equivalently, in a transaction vector w and use these elements in goal analysis, where the variables correspond to subgoals whose values are to be determined. We shall develop this around a simple example taken from Charnes, Cooper, and Ijiri, 1962. Then, in the next section we shall analyze some further implications of planning through a spread sheet, or simply, "spread sheet planning."

1. A Single Goal Problem

Let us use the beginning balance given in Table 1, Example 1 above. Let us, in addition, assume the following conditions on the firm's operations:

1) The firm can sell Materials simply by packing them by Machine 1 (processing rate: 2 hours per unit of Materials) or the firm can convert Materials into Finished Goods by Machine 2 (processing rate: 5 hours per unit of Finished Goods) and then sell the products after packing them by Machine 1 (processing rate: 3 hours per unit of Finished Goods.)

2) The available machine operating hours are 12 hours for Machine 1 and 10 hours for Machine 2.

3) The cost of Materials is $1 per unit, and from one unit of Materials we produce one unit of Finished Goods.

4) All processing costs and operating expenses are assumed to be fixed at \$2.5 per period regardless of the level of operation.

5) At the end of the period, the firm wants to maintain the inventory levels of Materials and Finished Goods at exactly the same level as at the beginning of the period.[1]

---

[1] I.e., this may be viewed as a generalization of breakeven (equilibrium) ideas from the activity (flow) accounts to the balance sheet (stock) accounts.

---

6) The firm can sell all Materials and Finished Goods packed through Machine 1 at the selling price of \$2 and \$1.5 per unit of Finished Goods and Materials, respectively.

7) All sales, purchases, and payments of expenses are made in cash.

8) The management goal is set at zero profit level (a breakeven point) for this analysis.

Let us first prepare a 4 by 4 spread sheet as in Example 1 above from the four accounts, Cash, Finished Goods, Materials, and Equity, denoted by C, G, M, and E, respectively.

Table V-2

A Spread Sheet

Dr. \ Cr.	C	G	M	E
C	*	(1)	(2)	(3)
G	(4)	*	(5)	(6)
M	(7)	(8)	*	(9)
E	(10)	(11)	(12)	*

Taking into account the conditions of the firm's transactions, we can give the following interpretation for each element in the spread sheet, which will be refered to as spread sheet variables. The entries in the diagonal cells (marked by *) represent intra-account

transfers, such as bank transfers in the case of Cash account, and play the role of "control entries" when sub-ledgers are prepared for an account. For our current analysis, however, the diagonal cells do not have significance and are therefore omitted from further consideration. The following interpretation is not unique but it is nevertheless concrete and can give us enough information about the nature of the transactions that the variables represent.

Table V-3
Spread Sheet Variables

No.*	Dr.	Cr.	
(1)	C	G	Cost of goods sold
(2)	C	M	Cost of materials sold
(3)	C	E	Gross profit on sales
(4)	G	C	Purchases of finished goods
(5)	G	M	Material costs
(6)	G	E	Other manufacturing costs
(7)	M	C	Purchases of materials
(8)	M	G	Finished goods spoilage returned as materials
(9)	M	E	Expenses chargeable to materials (e.g. freight)
(10)	E	C	Manufacturing, selling and administrative expenses
(11)	E	G	Finished goods spoilage with no salvage value
(12)	E	M	Materials spoilage with no salvage value

* As coded in Table V-2.

Let the amount of each one of the above transactions be represented by a variable X and let each X be assigned two subscripts, the first one representing debit entry account and the second one a credit entry account. For example, any transaction under (1) above would be represented by X_{CG}, while any transaction (2) would be

represented by X_{CM} and so on. Note that these notaions represent exactly the same factors which were represented by w_{ij} in the previous section. However, since, as we shall see in the next section, these variables act as a substitute for the subgoal variables x in the analysis of goals, we have used a capital X with two letter subscripts to denote the spread sheet variables, in order to relate both of these aspects in our analysis.

Now we refer to Table V-2 and immediately preceding technological and managerial conditions in order to develop constraints that are to be imposed upon the transactions. By the nature of the constraints, transactions (4), (6), (8), (9), (11), and (12) are not involved in any constraints, and hence in our X_{ij} notation.

(5.13a)	$3X_{CG} + 2X_{CM} \leq 12$	Machine hours Constraints (Instructions 1) and 2))
(5.13b)	$5X_{GM} \leq 10$	
(5.13c)	$X_{CG} = X_{GM}$	Balance Constraint (Instructions 5), 6), and the non-negativity condition on cash balance)
(5.13d)	$X_{GM} + X_{CM} = X_{MC}$	
(5.13e)	$-X_{CM} - X_{CG} - X_{CE} + X_{MC} + X_{EC} \leq 30$	
(5.13f)	$X_{CG} + .5X_{CM} = X_{CE}$	Gross Profit (Instruction 1)).
(5.13g)	$X_{EC} = 2.5$	Fixed costs (Instruction 4))
(5.13h)	$X_{CE} = X_{EC}$	Breakeven Constraint (Instruction 8)).

where all the variables are constrained to be non-negative.

These six variables X_{CM}, X_{CG}, X_{CE}, X_{MC}, X_{EC}, and X_{GM} in the above problem may be represented in an accounting network as follows:

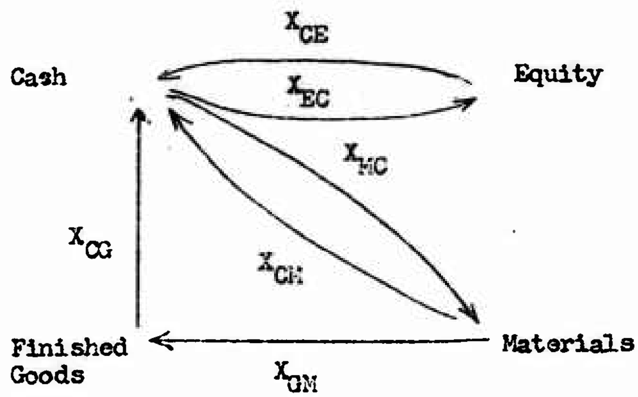


Figure V-2

Accounting Network II

Let $\check{Y}_C$, $\check{Y}_G$, $\check{Y}_M$, and $\check{Y}_E$ be the beginning balance of Cash, Finished Goods, Materials, and Equity accounts, respectively, and let ΔY_C, ΔY_G, ΔY_M, and ΔY_E be their increment, respectively, during the period. Then, relationship between X's and Y's can be given as follows by means of an incidence matrix associated with the above network.

$$\begin{bmatrix} \Delta Y_C \\ \Delta Y_G \\ \Delta Y_M \\ \Delta Y_E \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_{CM} \\ X_{CG} \\ X_{CE} \\ X_{MC} \\ X_{EC} \\ X_{GM} \end{bmatrix} \tag{5.13i}$$

Since by the instruction (4), (5), (6), (8), and the non-negativity condition on an ending cash balance we must have.

$$\text{(5.13j)} \qquad \begin{aligned} \Delta Y_C &\geq -30 \\ \Delta Y_G &= 0 \\ \Delta Y_M &= 0 \\ \Delta Y_E &= 0 \end{aligned}$$

The constraint (5.13c) - (5.13e) and (5.13h) were obtained by substituting the expression for Δ Y's given in (5.13i) into (5.13j).

For the goal analysis, we simplify (5.13a) through (5.13h) by appropriate substitutions and formulate the problem as follows:

$$\text{(5.14)} \qquad \begin{aligned} &\text{Minimize} && y^+ + y^- \\ &\text{subject to} && X_{CG} + .5X_{CM} - y^+ + y^- = 2.5 \\ & && 3X_{CG} + 2X_{CM} \leq 12 \\ & && 5X_{CG} \leq 10 \\ & && X_{CG},\ X_{CM},\ y^+,\ y^- \geq 0 \end{aligned}$$

Note that (5.14) is exactly the same as (3.3) in Chapter 3 except that there the variables X_{CG} and X_{CM} are replaced by x_1 and x_2. One of the solution to (5.14) is, therefore, $X_{CG} = 2$ and $X_{CM} = 1$, as derived in (3.3). By the equations (5.13c),(5.13d), and (5.13f), we get $X_{GM} = 2$, $X_{MC} = 3$, and $X_{CE} = 2.5$.

We now can prepare a spread sheet from this information. Here, however, we want to include balance sheet as well as flow information and for this we enlarge the ordinary spread sheet by adjoining one row and one column each for the beginning balances (BB) and for the ending balances (EB).

Table V-4

A Spread Sheet

Dr. \ Cr.	BB	C	G	M	E	EB	Total
BB					50		50
C	30		2	1	2.5		35.5
G	10			2			12
M	10	3					13
E		2.5				50	52.5
EB		30	10	10			50
Total	50	35.5	12	13	52.5	50	213

The figures in BB <u>column</u> -- 30, 10, and 10 -- represent the debit balances in the beginning balance sheet and the figure in BB <u>row</u> -- 50 -- represents the credit balance in the beginning balance sheet. See Table V-1. The entries in the double lined square result from solving the linear programming problem and substituting the results in (5.13a) - (5.13h). The figures in EB row and EB column are derived in the following manner. First we take the row sum and the column sum of each account. If the row sum exceeds the column sum for an account, the difference is recorded in EB row and if the column sum exceeds the row sum for an account, the difference is recorded in EB column. As a result, the row sum and the column sum for each account are equal.

It will be noted, as already discussed, that the entries in a row in a spread sheet all represent debit entries to that account. The entries in a column are credit entries, account by account. To make this completely clear, however, and to show the relationship to ordinary "T" accounts we may recapitulate the data on the transactions which effect cash account as

Table V-5

Cash Account

From Row C			From Column C	
Beginning Balance		$30.0	Purchases of Materials	$3
Sales			Expenses	2.5
Cost of Goods Sold	$2		Ending Balance	30
Cost of Materials Sold	1			
Gross Profit on Sales	2.5	5.5		
		$35.5		$35.5

The ending balance sheet is prepared from the entries in the EB row (debit balances) and EB column (credit balances) as follows:

Table V-6

Ending Balance Sheet

EB Row		EB Column	
Cash	$30	Equity	$50
Finished Goods	10		
Materials	10		
	$50		$50

This is exactly the same as the beginning balance sheet. See Table V-1. This result is due to the balance requirements for Finished Goods and Materials and the fact that the resulting transactions are required to breakeven.

The income statement for the period is prepared from the entries in E column (profit) and E row (loss) in the double lined square as follows:

Table V-7

Income Statement

E Row*		E Column*	
Expenses	2.5	Gross Profit	2.5
Net Profit	0		
	2.5		2.5

* In double-lined portion only of Table V-4.

2. A Multiple Goal Problem

Even though the linear programming formulation (5.14) takes into account only one goal (breakeven), the analysis can readily be extended to multiple goals as developed in the earlier chapters. We shall show only one of several possibilities for incorporating multiple goals in the problem by the following example.[1]

[1]Refer to Charnes, Cooper and Ijiri, 1962, p. 25 ff.

Suppose that management wants to attain a profit level of $2 after covering the fixed costs of $2.5. Suppose further that management is willing to deplete inventory, if necessary, to achieve the above objective but in this case management wants to deplete Finished Goods only after Materials are all depleted. Suppose, also, that management is willing to decrease cash balance, if necessary, only if both Materials and Finished Goods are completely depleted. Finally, suppose that so far as the desired profit level is attained and the cash and inventory levels are maintained at the levels of the beginning balances, this management does not care

about any overachievement of these levels.

Then, we have four goals which are arranged in an order of preemptive priority: (1) Achieve the level of contribution to overhead and profit at \$4.5. (2) Maintain a cash balance at least as great as \$30. (3) Maintain a balance of Finished Goods at least as great as \$10. (4) Maintain a balance of Materials at least as great as at \$10.

Assuming that the beginning balance of Finished Goods are already packed and ready for sale, but that the beginning balance of Materials are not yet packed (since they may be used to produce Finished Goods), the problem is formulated, after being simplified by appropriate substitutions as in (5.14), as follows:

(5.15)

$$
\begin{array}{llr}
\text{Minimize} & M_4 y_1^- + M_3 y_2^- + M_2 y_3^- + M_1 y_4^- & \\
\text{subject to} & & \\
(5.15a) & .5X_{CM} + X_{CG} - y_1^+ + y_1^- & = 4.5 \\
(5.15b) & 1.5X_{CM} + 2X_{CG} - X_{MC} - y_2^+ + y_2^- & = 2.5 \\
(5.15c) & X_{GM} - X_{CG} - y_3^+ + y_3^- & = 0 \\
(5.15d) & -X_{GM} - X_{CM} + X_{MC} - y_4^+ + y_4^- & = 0 \\
(5.15e) & 3X_{GM} + 2X_{CM} & \leq 12 \\
(5.15f) & 5X_{GM} & \leq 10 \\
(5.15g) & y_2^- & \leq 30 \\
(5.15h) & y_3^- & \leq 10 \\
(5.15i) & y_4^- & \leq 10 \\
 & X_{GM}, X_{CM}, X_{CG}, X_{MC}, y_1^+, y_1^-, y_2^+, y_2^-, y_3^+, y_3^-, y_4^+, y_4^- & \geq 0
\end{array}
$$

The constraints (5.15a) - (5.15d) represent the management's four goals. The constraint (5.15a), to achieve the contribution level of \$4.5, is the most important goal for the management, and, therefore, the highest preemptive pripority factor, M_4, is attached to the slack variable, y_1^-, which shows the under achievement of the goal.

The constraint (5.15b) takes care of the second goal, to maintain cash balance at least as great as \$30. The total cash receipts are given by $1.5X_{CM} + 2X_{CG}$ whereas the total expenditures are given by $X_{MC} + X_{EC} = X_{MC} + 2.5$. In order to maintain the cash balance at \$30 which is equal to the beginning cash balance, we must have

$$1.5X_{CM} + 2X_{CG} \geq X_{MC} + 2.5,$$

hence any shortage in the cash balance is represented by y_2^- on which the second highest priority factor M_3 is attached.

In addition, such a cash shortage compared with the minimum cash balance of $30 cannot exceed $30 since a negative cash balance is not allowed. This is represented as an environmental constraint, not as a goal, by (5.15g).

Similarly the third goal is represented by (5.15c) together with constraint (5.15h) to avoid a negative inventory balance, and the fourth goal by (5.15d) with the constraint (5.15i) by the same reason.

Note that y_i^+'s (i = 1, 2, 3, 4,) do not appear in the functional. This means that the management do not want to penalize any overachievement of the four goals.

The constraint (5.15e) and (5.15f) are the machine operating time constraints which we have discussed on various occasion.

A solution to (5.15) is as follows: $X_{GM} = 2$, $X_{CM} = 3$, $X_{CG} = 3$, $X_{MC} = 5$. The total contribution to overhead and profit, i.e., X_{CE}, is 4.5. Since all the operating expenses, X_{EC}, are fixed at $2.5, the net income for the period is $2 as desired. The ending cash balance is $33, which is above the beginning balance. However, the ending inventory level of Finished Goods was depleted from $10 to $9, whereas the ending inventory level of Materials are restored at the level of the beginning balance, $10. This means that there is no solution which results in the level of Finished Goods inventory being closer to $10 if we want to attain the two other goals which have a priority to this goal of restoring Finished Goods inventory. A spread sheet, a balance sheet, and an income statement prepared from this solution appear as follows:

Table V-8

A Spread Sheet

Dr. \ Cr.	BB	C	G	M	E	EB	Total
BB					50		50
C	30		3	3	4.5		40.5
G	10			2			12
M	10	5					15
E		2.5				52	54.5
EB		33	9	10			52
Total	50	40.5	12	15	54.5	52	224

Table V-9

Ending Balance Sheet

Cash	$33	Equity	$52
Finished Goods	9		
Materials	10		
	$52		$52

Table V-10

Income Statement

Expenses	$ 2.5	Gross Profit	$4.5
Net Profit	2		
	$4.5		$4.5

We have now translated a set of given goal into a set of subgoals and related them to standard accounting statements all by means of a spread sheet. From the solution of our linear programming model, we know that the multiple goals are not satisfied without violating the environmental constraints when the subgoals, represented by the variables X_{GM}, X_{CM}, X_{CG}, X_{MC}, X_{CE}, take the values of 2, 3, 3, 5, 4.5, and 2.5, respectively. By checking the alternate optimum solution, we know that X_{MC} may take any value between 5 and 8 but all other subgoal variables have to be exactly at the above specified levels. From these values for the subgoals, we can make a plan of operations for the coming period, since each one of the subgoal variables implies a particular type of operation, and the target value can be set for each type of operations.

On the other hand, once the planning stage is over, these spread sheet variables for the subgoals also serve for the management as a control device--i.e. as indicators. For example, in order to satisfy the multiple objectives, we know that total sales has to be \$10.5 (= $X_{CM} + X_{CG} + X_{CE}$). Therefore, by watching the accumlated total sales to date the management can judge whether or not all of the multiple goals are being attained. Furthermore, by means of a spread sheet, or an incidence matrix, described in the previous section, these indicators, which are concerned with transactions, can be converted into asset-equity balances. For example, from the above spread sheet we know that when the multiple goals are satisfied, the balance of Equity account should be \$52 and the balance of

Finished Goods account, $9, whereas Cash account should have a balance between $30 and $33 and Materials account between $10 and $13. Therefore, by our indicators this management can also assess whether or not the multiple goals are being attained on balance sheet accounts.

In the next chapter, these points will be more fully developed by means of an empirical application of this "spread sheet planning," for consolidating double-entry bookkeeping and linear programming. However, before going into any further development of "spread sheet planning," let us analyze in the next section some of the characteristics of spread sheet planning from the viewpoint of the goal analyses developed in the previous chapters and let us also do this together with the topic of aggregation of business activities as they are reflected in the double-entry bookkeeping system.

5. The Aggregation of Activities

The characteristics of such a spread sheet planning may be analyzed as follows, in connection with the problems of aggregation of activities.

Let us consider x as a vector whose components each represent a subgoal activity that takes place during an accounting period. Of course, the dimension of x may be astronomical for a large firm but this need not concern us here other than to note that then the accountant's job is to review each one of these activities, evaluate them, place them in appropriate positions in a spread sheet, or equivalently in a transaction vector w, and, finally, aggregate them to determine the aggregated values for each one of the components of w.

For our purpose, we may consider this whole process of accounting to be represented by an aggregation matrix C, through which x and w are related by $Cx = w$. The coefficients of the matrix C may be considered as dealing with the valuation and classification of the activities. For example, if the w vector is as given in Example 2 in Section 3, the first element in w, i.e., w_{31}, represents a debit entry to the material account and a credit entry to cash account. Then the first row of the matrix C has non-zero values only at columns which correspond to the material purchase (in cash) activities.

Then, the asset-equity balances for the purpose of preparing

financial statements may be presented as

(5.16) $\quad \Delta u = Tw = TCx$

where Δu is a vector of the amount of changes in the asset-equity balances and T is an incidence matrix discussed in Section 3.

On the other hand, these numerous activities, represented by the components of the vector x, collectively achieve, or come close to achieving, a set of goals b, the relationship being given by

(4.1) $\quad Ax = b$

as discussed in the previous chapter. We should like, if possible, to translate the given goals b into a set of more operational subgoals x, but if the dimension of x is very large, it is practically impossible to do so. We want, therefore, to have some ways to plan using aggregate variables, i.e., to translate the goals into some aggregates of the subgoal activities given by the components of the vector x. The spread sheet planning discussed in the previous section is a way of achieving this by replacing x with w, a vector for spread sheet variables which has fewer dimension than the vector x.

Since x and w are related by,

(4.2) $\quad Cx = w$,

we may express x by

(4.3) $\quad x = C^{+}w + C^{o}z \quad (z \in E^{n-r})$.

We have already examined in Chapter IV the result of replacing x in the goal-subgoal relationship (4.1) by the above expression (4.3), and also the condition on the C matrix in order to allow unique determination of the goal variables v from the spread sheet variables w. I.e., since (4.1) may be rewritten as

$$(4.4) \qquad v = AC^{+}w + AC^{o}z \qquad (z \in E^{n-r}),$$

the condition on C is, by Theorem 1 in Chapter IV, to be able to express each row of A by a linear combination of rows of C. We may interpret this condition as one which requires the accounting methods of valuation and aggregation of activities to be "in line" with the contributions of activities toward the goals of the management.

The same relationship can be derived for the constraints on the subgoal activities x, as given by

$$(3.1) \qquad Bx \leq h.$$

The constraints (3.1) may be replaced by

$$(5.17) \qquad BC^{\dagger}w + BC^{o}z \leq h,$$

by means of the expression (4.3). This implies that if each row of B can be expressed by a linear combination of rows of C the constraints may be converted into constraints on the variable w only, i.e.,

$$(5.18) \qquad BC^{\dagger}w \leq h.$$

While such a "perfect" consolidation[1] of x by means of the

[1]See also the discussion on aggregation and consolidation in Dorfman, Samelson, and Solow, 1958 and Rosenblatt, 1956.

spread sheet variables w is desirable, it may be practically impossible to have such a "perfect" consolidation. Perhaps, all that can be done is for accountants to consolidate "as perfectly as possible" for the goals which are at issue.[1]

[1]See, e.g., Charnes and Cooper, 1961, Appendix E.

The choice of aggregations and consolidations is not a technical matter only. The choice of variables as aggregates for the subgoal variables x also depends upon the goals of the management and the constraints imposed upon the subgoal variables, as well as psychological requirements of management. Therefore, whether or not spread sheet variables are good aggregates must be judged relative to ~~each~~ individual enviroment. Nevertheless, particularly when the goals are related with profits or funds flow, the spread sheet variables will perhaps show their greatest advantage as aggregates, since they are then also prepared as a result of the accounting theories and practices which have been developed in response to managerial concern with the alternate effects upon income determination.

In the next chapter, we shall report on an empirical application of spread sheet planning to a firm in the business game at Carnegie Institute of Technology as a means of illustrating the power derived from combining accounting and mathematical techniques.

CHAPTER VI

AN APPLICATION OF "SPREAD SHEET PLANNING"

1. Introduction

This chapter is concerned with an experimental application of our spread sheet approach to the analysis of goals developed in the previous chapter with specific reference to its use for (i) planning operations and (ii) evaluating various managerial policies, physical constraints, etc. In particular we want to show how to exploit our previous suggestions for using the more readily available categories of accounting costs without losing access to the corresponding, but more complex (and difficult to determine) categories of opportunity costs.[1] For this we shall draw upon a previous

[1] See, e.g., the discussion by Gould, 1962.

paper[1] which was developed around actual plays of a business

[1] This experiment was originally reported in a paper by Y. Ijiri, F.K. Levy, and R.C. Lyon, "A Linear Programing Model for Budgeting and Financial Planning, "ONR Research Memorandum No. 94, July, 1962 which is to be published in Journal of Accounting Research, Vol. I. A part of Section 2, 3, and 4 of this chapter was directly taken from this paper.

game -- the so-called Carnegie Tech game.

The Carnegie Tech Management Game is modeled after a "firm" in the packaged detergent industry. It offers the advantage of being immediately tractable without losing a rich array of business

features such as production, marketing, finance, etc.[1]

[1]The degree of complication of this game may be seen from the fact that the number of specific decisions each team makes during each month played ranges roughly from one hundred to four hundred. For further discussion of game details, see. e.g., Cohen, et. al., 1960, or Dill, et. al., 1961.

An application of only one of the several variations of the goal analyses developed in Chapter III is discussed in this chapter. However, this experimental model can easily be modified to meet these variations in Chapter III without changing the main structure of the model, since once we have succeeded in reflecting the operations in the real world on a mathematical model, the type of variations in the model such as ones discussed in Chapter III involve only mathematical manipulations of the model.

2. Synthesizing the Model

1. Accounting Period of Planning

The game involved here is a complex one; hence an adequate model of the kind with which we are concerned to examine will also be complex. Thus, to concentrate on points that are essential to "spread sheet planning" it is desirable to confine our model to only one period (a month) in the operations for one firm. This "month" is actually a "succeeding period" following on the completion of a particular round of decisions. But, of course, a sequence of one-month decisions do not add up, in general, to an overall desirable set of decisions. The latter would require either a full dynamic model or else the introduction of certain "safeguards" which will be examined when we discuss the implication of extending the model to multiple periods in a later section.

2. Goals

The first task in the construction of the model is to define clearly the management objectives. There are numerous issues here, but to save space we will not enter into them here and simply assume that the objective "maximization of net addition to retained earnings (after taxes and dividends)" is wholly adequate. We can, of course, take into account multiple objectives by means of the methods discussed in Chapter III, and other possibilities are also present such as attaining a "satisfactory profit" —, e.g., a profit level at

least as great as the previous year's profit, but for the moment we assume that maximization of net addition to retained earnings after taxes and dividends provides a "reasonable" objective for this management.[1]

[1]Of course, this objective will normally include many others as "equivalent objectives" in the sense that any one of them will lead to exactly the same optimum program. Also, the possibility of opportunity cost evaluations may produce still additional equivalences between multiple objectives when it is necessary to alter the constraints, etc. See, e.g., Charnes and Cooper, 1963.

3. Subgoals

We have generally assumed a divisionalized firm in our earlier analyses of goals, which we wanted to break down into more operational subgoals, and so we also assume that here. That is, given the indicated "overall goal," we next proceed to erect our model around subgoals whose values will collectively contribute towards the attainment of the goal. In this case the subgoals are represented in terms of the sixteen spread sheet variables listed below.

Table VI-1

Spread Sheet Variables

X_{CB}: Sales of government securities
X_{CR}: Collection of accounts receivable
X_{CL}: Borrowing
X_{CE}: Interest on government securities

X_{BC}: Purchases of government securities
X_{RG}: Cost of sales at standard cost
X_{RE}: Gross profit on sales
X_{GM}: Material costs

X_{GE}: Variable conversion costs
X_{MP}: Purchases of materials
X_{DC}: Payments of dividends and taxes
X_{PC}: Payments of accounts payable

X_{LC}: Repayment of loans
X_{EF}: Depreciation
X_{EP}: Manufacturing and operating expenses
X_{ED}: Accruals of income taxes and dividends

As mentioned in the previous chapter, the letter pair subscripts in the above spread sheet variables indicate their effects on balance sheet accounts. Thus, using "X" to represent any "amount," we have "X_{ij}" to represent the corresponding debit to account "i" and the credit to account "j."

In our case the accounts i and j are separately "named" as follows:

Table VI-2

Balance Sheet Accounts

Asset Accounts

C: Cash

B: Government securities

R: Accounts receivable

G: Finished goods

M: Raw materials

F: Fixed assets

Equity Accounts

P: Accounts payable

D: Dividends and taxes payable

L: Loans payable

E: Stockholders' account

In this model, we eliminated all income statement accounts and used Stockholders' account (E) whenever debit or credit entries are required to an income statement account. We have already examined this in the previous chapter. Even though we need to decompose the stockholders' account (or capital account) for more complicated situations, in this experiment we can thus greatly ~~reduced~~ the size of the problem without affecting its essential nature.

4. Beginning Balance Sheet

For our purposes we now produce the following balance sheet from this firm's records and key the amounts to their corresponding symbols as in Table 3, below.

Table VI-3

Beginning Balance Sheet

C:	Cash	$ 7,260,000	P:	Accounts Payable	$3,592,000
B:	Government Securities	12,000,000	D:	Dividends and Taxes Payable	2,922,000
R:	Accounts Receivables	6,999,000	L:	Loans Payable	4,400,000
G:	Finished Goods	4,032,000	E:	Stockholders' Account	46,876,000
M:	Raw Materials	1,499,000			
F:	Fixed Assets	26,000,000			
		$ 57,790,000			$57,790,000

Next we introduce the following convention. We use the one symbol K with an appropriate subscript to represent any of these figures in the beginning balance sheet. For instance, K_C (= $7,260,000) refers to the opening cash balance, while K_B (= $12,000,000) refers to the opening balance of the government securities account, and so on.

5. Constraints

We now impose upon the sixteen subgoal variables a set of constraints. These constraints, which are described below, mostly come from the enviroment of the firm in the business game, especially

from the rules of the game. Here we proceeded by aggregating (or consolidating) all of the firm's products into an equivalent single standard product. This, of course, raises "information content problems," but we shall later discuss this topic in more detail. For the moment we assume that all aspects of this aggregation -- e.g., in inventory account --are wholly satisfactory. That is, we assume that this standard product yields a categorization that is pertinent to both our subgoals and our overall objectives or goals.

Since all the subgoal variables are to be valued in dollar terms, the constaints are also stated in terms of dollars. In order to do this, we set up standard costs by means of which the constraints are all converted into dollar terms. Since we are in the planning stage, we do not immediately worry about the variances between the actual and planned program values which later enter into the problems of control. Thus proceeding on this line and utilizing actual game data, the selling price and cost factors, per unit of the standard product, are estimated to be as follows:

Table VI-4

Selling Price and Standard Cost

Selling price	$10.00
Standard cost of finished goods	
Material costs	
Standard material consumption:	
1 unit per unit of finished goods	
Standard material purchase cost: $1/unit	$1.00
Variable conversion costs	1.10
Standard cost of finished goods	$2.10
Standard purchase cost of materials	$1.00

Similar valuation problems arise on other accounts, too — *viz.*, government securities, accounts receivable, and fixed assets. Here, however, the rules of this particular game may be utilized to produce resulting simplifications that are valid for our purposes. For example, under these rules, government bonds are the only securities that a firm can hold, and all purchases and sales are made only at face value plus or minus a known adjustment for interest accrued. This means, of course, that an accurate valuation can always be effected by using the face value of these bonds and accomodating any adjustment by means of the interest income account. Similarly, no bad debts appear in this game, and so no allowance need be made for a valuation reserve; consequently accounts receivable amounts are always at their face value. Also depreciation of fixed assets is determined by the declining balance method and the fixed

asset balances are determined by deducting the resulting reserve to arrive at an undepreciated asset amount.

Therefore, the valuation problems of assets other than inventories are greatly simplified in this model. However, this is because we wanted to take maximum advantage of our knowledge concerning the environment of the firm in the business game,

in order to represent the firm's operations and environment to a satisfactory degree by a simplified model. This does not mean that we cannot handle more complicated cases. Not only can "standard cost methods" such as those used for inventory valuation be but also further alterations can be introduced such as asset liquidation, stretching of accounts vs. borrowing, etc., when these, too, need to be actually considered in more complicated cases.[1]

[1]See, e.g., Charnes, Cooper, and Miller, 1959

Allowing for the above aggregation and valuation considerations, we now prepare nineteen constraints which are to be imposed upon the sixteen subgoal variables listed in Table 1. Although some of these will now be developed separately, it is to be emphasized that this is merely for convenience of exposition. Although most of these constraints come from the environment of the firm in the business game and the rules of the game, the validity and reasonableness of these constraints must not be evaluated independently. Stating the the matter differently, all constraints and objectives must generally

be accorded joint consideration if a valid model is to be achieved.

We now synthesize these sixteen constraints in the light of the following considerations.

(1) The firm's sale of government securities cannot exceed the beginning balance in this account--viz.,

(6.1) $$X_{CB} \leq K_B$$

(2) Interest on government securities, as earned, is received in cash for the government securities on hand at the <u>end</u> of the period. The monthly interest rate is 0.229%. Hence

(6.2) $$X_{CE} = .00229\ (K_B + X_{BC} - X_{CB})$$

To see what this expression means note that K_B is the opening balance for government securities while X_{BC} refers to government securities purchased during the period, by offsets to cash. Similarly, X_{CB} refers to the cash debits realized by sales of government securities during the period. Hence $(K_B + X_{BC} - X_{CB})$ correctly states the ending balance when, as is true under the conditions of the Carnegie Tech game, all such transactions are affected only in cash. Thus applying the indicated interest rate of 0.229%, we obtain exactly X_{CE} the amount which is debited to cash and credited to the stockholder's account.[1]

[1]In the Carnegie Tech Game, all interest income is realized as immediate cash income, and is the only cash income in the game. Also, the price of the bond is held constant throughout the Game periods as noted before.

(3) Since the firm's terms of sale are 30 days net, its maximum collection of receivables during the month is given by the beginning balance in the receivables account.

(6.3) $$X_{CR} \leq K_R$$

(4) The beginning-of-the-period cash balance limits the purchase of securities.

(6.4) $$X_{BC} \leq K_C$$

(5) Since the firm's selling price during the period studied is \$10.00/unit, while its standard cost of production is \$2.10/unit, the following constraint represents the gross profit on a unit sale in terms of the corresponding deduction from finished goods inventory:

(6.5) $$X_{RE} = \frac{10 - 2.1}{2.1} X_{RG}$$

or $$X_{RE} = 3.76 X_{RG}$$

(6) Standard cost of finished goods (\$2.10/unit) includes material costs (X_{GM} = \$1/unit) and variable conversion costs (X_{GE} = \$1.10/unit), which consists of direct labor cost and direct overhead.

(6.6) $$X_{GE} = 1.1 X_{GM}$$

(7) Production capacity limits conversion during the period. Here production capacity is expressed in terms of the value of raw materials (at standard cost) that can be converted.

(6.7) $$X_{GM} \leq 1,300,000$$

(8) Conversion is limited to raw materials on hand at the beginning of the month.[1]

(6.8) $X_{GM} \leq K_M$

---

[1]In this case K_M = \$1,499,000 so that the preceding constraint, (6.7), supersedes this constraint. We have, however, represented both constraints for the sake of completeness in cases where K_M < \$1,300,000.

---

(9) Market conditions limit sales in the coming month to 2,000,000 units[1] (\$4,200,000 at standard cost).

(6.9) $X_{RG} \leq 4,200,000$

---

[1]Obtained from the relevant market forecast which under this firm's approach produces these figures: an expected rate, a maximum and a minimum. The maximum is entered here as an exact upper limit for subsequent study by reference to the relevant "dual evaluators" which are to be used as "opportunity cost" figures for assessing possible sales improvements and the expenditures needed to achieve them etc. Refer to analyses in Section 5 of this chapter.

---

(10) Sales during the month are further limited to the completed units on hand at the beginning of the month.[1]

(6.10) $X_{RG} \leq K_G$

---

[1]Cf. Footnote 1, in (9). It should also be observed that this constraint is intended to reflect the lag on shipments from the plant to the relevant distribution centers.

---

(11) Repayment of loans is limited by the outstanding loan balance at the beginning of the month.

(6.11) $X_{LC} \leq K_L$

(12) Because the firm is allowed 30 days for payment of materials purchased on credit, it has adopted a general policy of not paying the accruals incurred on accounts payable that result

from material purchases during the month. Thus, the payment of accounts payable is limited to the sum of the beginning balance of accounts payable plus expenses accrued during the month, excluding payables accrued from material purchases.

(6.12) $$X_{PC} \leq K_P + X_{EP}$$

(13) Monthly depreciation charges are .833 per cent of the net fixed assets owned at the beginning of the month.

(6.13) $$X_{EF} = .00833\ K_F$$

(14) Expenses to be incurred during the coming period, including both manufacturing cost (other than material cost) and operating expenses, consist of the following four items: (a) Fixed operating expenses, $2,675,000; (b) Variable conversion cost, X_{GE}; (c) Effective interest penalty for discounts not taken on accounts payable (3.09 per cent per month); (d) Interest (monthly rate, .291 per cent) on loans payable at the end of the month. Thus, putting these all together,

(6.14)

$$X_{EP} = 2{,}675{,}000 + X_{GE} + .0309\ (K_P + X_{EP} - X_{PC}) + .00291\ (K_L + X_{CL} - X_{LC})$$

Note that X_{EP} includes all expenses (other than material costs) regardless of whether they are chargeable to finished goods. This means that the stockholders' account is understated by the amount equal to X_{GE} since this portion represents costs applicable to future periods by means of charges to inventory. However, this is adjusted by X_{GE} which adds back to the stockholders' account precisely the amount to be charged to inventory. This manipulation was necessary

to take into account the interest charges which are applicable to all X_{EP} including X_{GE}.

(15) Income tax is accrued at 52 per cent of the net profit. In addition, the firm's policy is to declare a dividend equal to \$83,000 plus (minus) 5% of the excess (shortage) of what it considers a standard net profit after taxes, \$1.860,000.

$$(6.15) \quad X_{ED} = .52\,(X_{CE} + X_{RE} + X_{GE} - X_{EF} - X_{EP}) + 83,000$$
$$+ .05\,[.48\,(X_{CE} + X_{RE} + X_{GE} - X_{EF} - X_{EP}) - 1,860,000]$$

Simplifying,

$$(6.15) \quad X_{ED} = .544X_{CE} + .544X_{RE} + .544X_{GE} - .544X_{EF} - .544X_{EP} - 10,000$$

(16) Company policy is to maintain a minimum cash balance of \$4,000,000 at the end of each period.

$$(6.16) \quad K_C + X_{CB} + X_{CR} + X_{CL} + X_{CE} - X_{BC} - X_{PC} - X_{DC} - X_{LC} \geq 4,000,000$$

(17) Because of an impending price rise in the following period -- i.e., in the period following the one under study -- the firm requires that the end-of-the-month finished goods inventory be at least as great as the minimum sales expected during the next month (\$3,570,000 at \$2.10 standard cost/unit).[1]

[1] This is one way in which a one-period model can be lashed into succeeding periods. Note, for instance, that this amount, too, is subject to treatments by means of the dual evaluators as we shall see in Section 5. Such an analysis will then show how much is being lost from current profit opportunities in order to provide for the future in this manner and the firm's management can then alter this figure or request an extension of the model to explore the consequences of different alterations, etc.

$$(6.17) \quad K_G + X_{GM} + X_{GE} - X_{RG} \geq 3,570,000.$$

(18) The firm expects to produce 1,200,000 units in the succeeding month. Thus the raw materials on hand at the end of the month must be sufficient for this production. (At \$1.00 raw materials cost/unit, the end of the month raw materials balance must be \$1,200 ,000.)[1]

(6.18) $$K_M + X_{MP} - X_{GM} \geq 1,200,000$$

---

[1]See preceding footnote.

---

(19) All outstanding income taxes payable and dividends declared, including those accrued or declared during the month, have to be paid by the end of the coming month.

(6.19) $$K_D + X_{ED} - X_{DC} = 0$$

In addition to these constraints, all subgoal variables (X's) are required to be non-negative. This is done to avoid any confusion between the debit and credit entries which would otherwise have to be given a reverse (credit-debit) interpretation for negative values of any X.[1] Of course, the non-negativity conditions can

---

[1]See the network analysis in the previous chapter.

---

be eliminated but for the present purposes, it simplifies the exposition to arrange the model so that

all variables are required to be non-negative.

6. The Formulation of the Linear Programming Model

The given management objective of maximizing net additions to retained earnings (ΔK_E) is now represented by a combination of subgoal variables as follows:

$$\Delta K_E = X_{CE} + X_{RE} + X_{GE} - X_{EF} - X_{EP} - X_{ED}. \tag{6.20}$$

Therefore, the full linear programming model is set forth as follows:

(6.20) Maximize $(\Delta K_E =)\ X_{CE} + X_{RE} + X_{GE} - X_{EF} - X_{EP} - X_{ED}$

subject to

(6.1)	X_{CB}	$\leq K_B$
(6.2)	$X_{CE} - .00229\ X_{BC} + .00229\ X_{CB}$	$= .00229\ K_B$
(6.3)	X_{CR}	$\leq K_R$
(6.4)	X_{BC}	$\leq K_C$
(6.5)	$X_{RE} - 3.76\ X_{RG}$	$= 0$
(6.6)	$X_{GE} - 1.1\ X_{GM}$	$= 0$
(6.7)	X_{GM}	$\leq 1{,}300{,}000$
(6.8)	X_{GM}	$\leq K_M$
(6.9)	X_{RG}	$\leq 4{,}200{,}000$
(6.10)	X_{RG}	$\leq K_G$
(6.11)	X_{LC}	$\leq K_L$
(6.12)	$X_{PC} - X_{EP}$	$\leq K_P$
(6.13)	X_{EF}	$= .00833\ K_F$

(6.14) $.9691\ X_{EP} - X_{GE} + .0309\ X_{PC} - .00291\ X_{CL} + .00291\ X_{LC}$

$$= 2{,}675{,}000 + .0309\ K_P + .00291\ K_L$$

(6.15) $X_{ED} - .544\ X_{CE} - .544 X_{RE} - .544\ X_{GE} + .544\ X_{EF} + .544\ X_{EP} = -10{,}000$

(6.16) $-X_{CB} - X_{CR} - X_{CL} - X_{CE} + X_{BC} + X_{PC} + X_{DC} + X_{LC} \leq K_C - 4{,}000{,}000$

(6.17)	$-X_{GM} - X_{GE} + X_{RG}$	$\leq K_G - 3{,}570{,}000$
(6.18)	$-X_{MP} + X_{GM}$	$\leq K_M - 1{,}200{,}000$
(6.19)	$-X_{ED} + X_{DC}$	$= K_D$

$X_{CB}, X_{CR}, X_{CL}, X_{CE}, X_{BC}, X_{RG}, X_{RE}, X_{GM},$

$X_{GE}, X_{MP}, X_{PC}, X_{DC}, X_{LC}, X_{EF}, X_{EP}, X_{ED} \geq 0$

That is, we are to search for amounts X_{ij} which satisfy all constraints, including non-negativity for all variables, and among the entire collection of such values we are to single out one set that makes the net addition to retained earnings a maximum. The solution methods of linear programming -- e.g., the simplex method[1] --

[1] We used this method as one part of the available G-20 electronic computer code without attempting to apply or devise other, possibly more efficient, methods. For further remarks on this topic see Charnes, Cooper and Ijiri, 1962.

are designed, of course, to produce exactly this result.

3. The Results

By substituting the figure in the beginning balance sheet (Table 3) for the K's in the above constraints and by applying the simplex method, a solution to the above problem was readily achieved. We now summarize the results in the following series of Tables. First in Table 5 we list the transactions which achieve the maximum possible net addition to retained earnings.

Table VI-5

Optimum Projected Transactions
(All figures are in dollar)

SB:	Sales of government securities	9,703,000
CR:	Collection of accounts receivable	6,999,000
CL:	Borrowing	-
CE:	Interest on government securities	6,000
BC:	Purchases of government securities	-
RG:	Cost of sales at standard cost	3,192,000
RE:	Gross profit on sales	12,002,000
GM:	Material costs	1,300,000
GE:	Variable conversion costs	1,430,000
MP:	Purchases of materials	1,001,000
PC:	Payments of accounts payable	7,697,000
DC:	Payments of dividends and taxes	7,871,000
LC:	Refund of loans	4,400,000
EF:	Depreciation	217,000
EP:	Manufacturing and operating expenses	4,105,000
ED:	Accruals of income taxes and dividends	4,949,000
	Net addition to retained earnings	4,167,000

Next, we utilize this information to obtain the following balance sheet and income statement projections for the period at issue.

Table VI-6

Projected Balance Sheet
(End of Period)

C:	Cash	$4,000,000	P:	Accounts Payable	$1,001,000
B:	Government Securities	2,297,000	D:	Dividends and Taxes Payable	-
R:	Accounts Receivable	15,194,000			
G:	Finished Goods	3,570,000	L:	Loans Payable	-
M:	Raw Materials	1,200,000	E:	Stockholders' Account	51,043,000
F:	Fixed Assets	25,783,000			
		$52,044,000			$52,044,000

Table VI-7

Projected Income Statement

Sales (RG + RE)		$15,194,000
Cost of goods sold (RG)		3,192,000
Gross profit (RE)		12,002,000
Manufacturing costs and operating expenses (EP)	$4,105,000	
Depreciation (EF)	217,000	
Manufacturing costs charged to finished goods (GE)	- 1,430,000	
Unabsorbed manufacturing costs and operating expenses		2,892,000
Operating profit		9,110,000
Interest income (CE)		6,000
Net income before tax		9,116,000
Accruals of income taxes and dividends (ED)		4,949,000
Net addition to retained earnings		$4,167,000

Notice that we now have the main documents that are usually deemed to be pertinent for planning purposes. Additional documents -- e.g., flow of funds statements, cash budgets, etc. -- can also be generated if desired.[1] It is not proposed, however, to pursue

[1] See Charnes, Cooper and Ijiri 1962, for further discussion.

this topic here. We note instead how the results of this analysis can be used for planning and decision making by various kinds of management.

The details of Table 2 can readily be translated into instructions for operating managers or other persons who are not immediately concerned with the overall aspects of planning. For instance, the production manager may be directly instructed to plan to produce \$2,730,000 worth of finished goods during this period and further to allocate \$1,300,000 of this amount for the purchase of raw materials and to expend the rest (\$1,430,000) on conversion costs. In turn, the purchasing department may be instructed to purchase \$1,001,000 worth of raw materials, and so on. This means that the management goal -- maximum addition to equity account, after taxes and dividends, under the indicated constraints -- is now translated into a set of more operational subgoals which are given to relevant organizational segments as goals imposed upon them.

Of course, the usual information for a coordinated assessment and review by top management is also available in the form of projected balance sheets, income statements, etc. In addition,

the linear programming solution provides valuable byproduct information in the form of the dual evaluators.[1] We propose therefore to elaborate on this topic in the next section and then show how this information can be assembled to form a new type of accounting document for use in integrated management planning.

---

[1]The G-20 code (like most linear programming codes) provides these data automatically, at the end of its run, as well as other (equally valuable) results which enable the user to test the program's sensitivity to errors in the data, etc. See, e.g., Charnes and Cooper, 1961, Ch. XIII for a discussion in the context of the so-called revised simplex code for IBM computers.

4. An Accounting Statement for the Dual Evaluators as Opportunity Costs

A dual evaluator, at least in our managerial-accounting-economics context, indicates the change in net addition to retained earnings that can be secured if the constraint corresponding to the given evaluator were relaxed by one dollar. For example, the dual evaluator of constraint number (6.7) which is a constraint on production capacity has a value of \$3.594936. This means that if production capacity were increased so that exactly one additional dollar's worth of raw material can be processed, then, retained earnings will be increased by \$3.594936. This figure, which is obtained (automatically) as a so-called "evaluator," really summarizes a whole complex of interrelated opportunities.[1] Hence some further accounting aids to managerial interpretations and uses are indicated.

[1] I.e., these are "mutatis mutandis" opportunity costs rather than of the "ceteris paribus" variety. The difference lies in the fact that the latter considers variation in one or a few variables with all other variables fixed, whereas the former assumes that all variables are optimally adjusted to the new opportunities that may be made available. See Charnes and Cooper, 1961, for further discussions.

In order to show that the dual evaluators take into account every constraint of the model in "mutatis mutandis" fashion, the evaluator of production capacity will now be analyzed in more detail. Therefore, suppose it is desired to alter the firm's raw material processing capacity by one dollar worth of materials, i.e. one unit, since the material cost is \$1/unit. This activates a whole complex of transactions which are set forth in the following statement.

Table VI-8

Analysis of Opportunity Costs*

(1)	Incremental Sales Resulting ($2.1 x 3.76)		$9.996000
	Less: Cost of Incremental Sales		2.100000
(2)	Incremental Gross Profit		7.896000
(3)	Deduct: Income Taxes and Dividends (54.4%)		4.295424
(4)	Net retained from sales after taxes and dividends		3.600576
(5)	Deduct: Opportunity Cost of the Cash Needed to Finance Expansion:		
	Cash Required for Raw Materials**	$ 0	
(5a)	Conversion to Finished Product	1.100000	
	Total	$1.100000	
(5b)	Add: Cash Required for Income Taxes and Dividends	4.295424	
(5c)	Total Cash Required	$5.395424	
(6)	Total of Interest Earnings Foregone on Securities Sold ($5.395425 x 0.00104533)		0.005640
(7)	Net Retained Realization per Unit Increase in Production Capacity		$3.594936

* This statement is based on optimal adjustments being effected in all pertinent transactions.

** Note: All required raw materials are available from inventory on hand.

Notice now that the end result is the $3.594936 figure that was predicted by the dual evaluator. Notice also that the indicated transactions are each carried out optimally (relative to one another) in order to produce this optimal return from the indicated one unit of increased capacity. The single figure of $3.594936 summarizes all of these transactions to the one net end effect, as indicated, but this obviously needs accounting elaboration if the purposes of managerial use and understanding are to be adequately served.

The above statement is intended to supply this kind of service.

The data of Table VI-8 were drawn from the byproduct results furnished automatically by the computer calculations after appropriate supplementation by economic-accounting considerations. We now elaborate on this as follows:

In (1) of Table VI-8 we observe that the incremental capacity increase has resulted in a sale of goods. The needed raw material is presently on hand, since X_M = \$1,499,000 (Table VI-3). is greater then X_{GM} = \$1,300,000 (Table VI-5). Therefore an additional raw material unit is drawn from inventory. This unit is processed at \$1.10 conversion of finished goods which is valued at standard costs, \$2.10. The unit of finished goods can be sold at once, since the firm has unfulfilled demand in the present month of \$1,008,000 of product (valued at standard cost -- see the constraint (6.9) remembering X_{RG} = \$3,192,000 in the optimum solution.) Segments (2), (3), and (4) give the net profit realized from this sale as determined by the constraints (6.5) and (6.15) of the model.

The transaction analysis is not yet complete, as the gross profit (\$3.600576) is higher than the dual evaluator (\$3.59436). The difference (\$.00564) can be accounted for by turning to the cash account and recalling that the program has puched up against the \$4,000,000 stipulated minimum balance. Therefore, the additional cash outlays needed for production must be obtained in one of three ways: (a) sell government securities, losing interest income at the monthly interest rate 0.229%; (b) borrow from bank at the monthly interest rate 0.291%; (c) postpone payments on accounts payable, paying interest at the monthly interest rate 3.09%. The least costly of these three ways is obviously achieved in the sale of securities. The \$.00564 shown opposite (6) in the above statement represents the opportunity cost incurred because of the interest income that must be foregone if securities are sold to obtain the needed cash.

This is derived as follows: (5a) and (5b) gives the amount of cash needed to increase production by one unit. (5a) represents the conversion cost which is paid at the end of the month. (5b) represents dividends and taxes payable on the profit generated in (4), and cash must be obtained to pay these at the end of the month by (6.19). The total

cash requirement is therefore the sum of (5a) and (5b), or $5.395474,[1] which is the figure shown opposite (5c). (6) multiplies **the number of dollars needed by the opportunity cost per dollar**

---

[1]Note that two other cash flows (accounts receivable and payments on raw material accounts) are affected by the capacity increase. However, they have no effect on the flow in the present month, as they are not received (or paid) until the following month.

---

to obtain the total opportunity cost of obtaining cash.

The opportunity cost per dollar is obtained via the following reasoning: The firm foregoes interest income of $.00229 for every dollar of securities sold. However, the loss is reduced to $.00104424 by taking into account savings on taxes and dividends, 54.4%x$.00229. Since the interest income would normally be collected in cash at the end of the month, the $.00104424 of income foregone reduces the end-of-month cash valance by $.00104424. This reduction in cash balance violates constraint (6.16) (minimum cash balance constraint) and to compensate for this reduction the firm must sell an additional $.00104424 of securities and forego (.00104424 x .00104424) dollars of net interest income, as well as (.00104424 x .00104424) reduction in the cash balance.

The total reduction of net profits after taxes and dividends resulting from this sale of securities needed to obtain an additional dollar of cash is given by the sum of the infinite series:

$$.00104424 + (.00104424)^2 + \ldots = \frac{.00104424}{1 - .00104424} = .00104533$$

Therefore .00104533 appears as the unit opportunity cost in (6).

Finally, in (7) we see that the net effect of relaxing the production capacity constraint by one unit is equal to the difference between net profit (after taxes and dividends) realized from the sale of goods and the opportunity cost of securing additional cash required. That is, $3.600576 - $.005640 = $3.594936. This figure is precisely equal to the dual evaluator associated with production capacity.

Notice, in particular, now that any one dual evaluator provides the opportunity cost information that is relevant when every possible transaction is altered to a new optimum level in response to a change anywhere in the system. This is, in fact, the intended meaning of the term "mutatis mutandis" which has been accorded

to this way of reckoning opportunity costs, in contrast to other approaches which assume that only a few variables are adjusted while all others are held constant.

The range in which the dual evaluator is valid is limited. For example, the increase in the production capacity contributes to the net addition to retained earnings (ΔK_E) by the amount of \$3.59 per unit increment in the production capacity only up to a given limit of such alteration. After passing that limit, the further expansion contributes to ΔK_E at a smaller rate and eventually the contribution vanishes at a certain point, beyond which the expansion of the production capacity has no effect on ΔK_E due to other constraints which make the production capacity constraint redundant. On the other hand, a <u>decrease</u> in production capacity results in a decrease in ΔK_E at the rate of \$3.59 per unit, but this rate <u>increases</u> as we decrease the production capacity beyond a certain critical point, and keeps increasing each time the production capacity passes a critical point in its decreasing process. Such critical points and changes in the rate can be calculated by the simplex method as we shall show in the next section.

5. Feedbacks for Management

A number of variations of the above analysis can be made based on the same model with minor modifications. Among them we shall describe two examples of the extensions of the analysis in the previous sections. One of them is an extension of the dual evaluator analysis in the previous section by means of the method described in Chapter IV and in Appendix C. Then, we shall show the second example of possible extensions of the model by means of a so-called "sensitivity analysis" of a linear programming problem. In contrast with a dual evaluator analysis which manipulates the stipulation level of a constraint, i.e., a constant in the right hand side of an inequality in (6.1) - (6.19), a sensitivity analysis -- or at least an extended sensitivity analysis[1] -- tries to manipulate a coefficient in a constraint, i.e., a coefficient attached to a variable X_{ij} in the left-hand side of a constraint, to see the effect on the objective function.

[1] See, e.g., Charnes and Cooper, 1963.

1. An Extension of a Dual Evaluator Analysis

Let us proceed to an extension of the dual evaluator analysis using an example. As mentioned in the last paragraph in the previous section, a dual evaluator associated with a constraint does not remain constant for all possible stipulation levels of the constraint. Take, for example, the constraint (6.9) which limits the sales by the market conditions, (I.e. the maximum sales volume we can attain in the coming month is estimated to be 2,000,000 units,

or \$4,200,000 at the standard cost of goods sold.) The dual evaluator associated with this constraint at the optimum solution given in Table 5 is zero. This means that increase or decrease in this sales estimate by one dollar does not affect the net addition to retained earnings.

Notice that X_{RG}, cost of sales at standard cost, is constrained to be less than or equal to \$4,032,000 by (6.10) which supersedes the constraint (6.9). Furthermore, by (6.17) X_{RG} plus the minimum inventory level at the end of the period, i.e., \$3,570,000, must together be greater than or equal to the production during the period (X_{GM} + X_{GE} or 2.1 X_{GM} by (6.6)) plus the beginning inventory, \$4,032,000. However, since the production volume (represented by X_{GM}) is limited to \$1,300,000 (at material cost) by (6.7), this puts a constraint on X_{RG} which requires it to be at most equal to \$3,192,000 ($X_{RG} + 3{,}570{,}000 \leq 2.1 \times 1{,}300{,}000 + 4{,}032{,}000$). This clearly supersedes (6.10) and of course (6.9) which we are interested in.

However, this redundant sales constraint (6.9) becomes an effective constraint if by a change in the market forcast the estimated maximum sales is reduced below \$3,192,000, making the production constraint (6.7) redundant. Then the dual evaluator associated with (6.9) becomes no longer equal to zero. It becomes equal to \$1.7118742 for \$1 increment in cost of sales at standard cost. Since the standard cost of the finished goods is \$2.1 per unit, this means that one unit increase or decrease in the sales

limit will bring \$3.54936 (= \$1.7118742 x 2.1) increase or decrease in the net addition to retained earnings, provided the current sales limit is less than \$3,192,000. It may be clear from Table VI-8 that this is precisely equal to the dual evaluator associated with the production constraint (6.7).

If the current sales limit is exactly \$3,192,000, the dual evaluator associated with this constraint is two-fold; i.e., for any increase in the sales limit, the net addition to retained earnings is not affected due to the production constraint (6.7) together with the minimum inventory balance constraint (6.17), whereas for any decrease in the sales limit the net addition to retained earnings is reduced by \$1.7118742 per \$1 decrease in the sales limit. That is, the dual evaluator is zero for an increment and \$1.7118742 for a decrement in the sales limit if the current sales limit is exactly at \$3,192,000.

This value of the dual evaluator remains constant for further reductions in the sales limit until \$462,000 is reached at which point the dual evaluator is increased to \$1.7124218. To see the reason for this change let us analyze the production side of the problem.

As we decrease the sales limit further, the production volume is also decreased since, as we shall see later, any surplus production causes a decrease in the net addition to retained earnings. Let us analyze how much the latter is decreased by one unit of surplus production, i.e. inventory at the end of the period over and above

the minimum inventory level imposed by (6.17). For one unit of surplus inventory, we have to purchase one dollar worth of materials (or reduce the material inventory by the same amount) and pay conversion costs of \$1.1. Neither is an expense chargeable to the current period. However, we have to consider the implications for interest. The material purchase cost need not be paid in this month since the firm has a 30 day credit. But the firm has to pay the conversion cost of \$1.1 by the end of the month if they want to avoid paying 3.09% interest on any overdue amount and this happens to be the most expensive way of financing. Here again the least expensive way of financing is to sell securities which when done, reduces the net addition to retained earnings by \$0.00104533 for each dollar sales of securities, or \$0.00114986 for \$1.1 sales of securities. This is the opportunity cost of producing one unit of product for inventory accumlation. Or stating the matter differently, this is the opportunity saving which will be obtained whenever we reduce the production volume by one unit if we set aside its effect on the marketing side of the problem.

As we decrease the sales limit, we, at the same time, decrease the production volume so that the ending inventory is always at a minimum (see (6.17). By doing this we can save the loss from decreased sales by \$0.00114986 per unit decrease in sales or \$0.00054755 per dollar decrease in cost of sales. This has already been taken into account in the dual evaluator for $462,000 \leq X_{RG} \leq 3,192,000$, i.e., \$1.7118742, since the dual evaluator is the result obtained when everything is adjusted optimally. Notice that at $X_{RG}=462,000$,

the optimum (or minimum) $X_{GM} = 0$. That is, the sales can be made entirely from the beginning inventory without violating the minimum inventory balance constraint. Since the production volume has already been pushed down to zero, the savings which have been obtained from reducing the production volume vanishes. This means that any further reduction in the sales limit will reduce the net addition to retained earnings by \$1.7124218 (= \$1.7118742 + \$0.0054755) per dollar decrease in cost of sales or by \$3.596086 (= \$3.594936 + \$0.00114986) per unit decrease in sales volume. The latter figure may be obtained directly from Table VI-8 by eliminating (5a).

In summary, a dual evaluator is a step function of a stipulation level as we have observed in the above specific example. We next proceed to indicate a managerial use of the dual evaluator thus explored for all possible ranges of the stipulation level, in the following way.

We know that at X_{RG} =\$3,192,000 the net addition to retained earnings (ΔK_E) is \$4,166,665. We also know that a one dollar decrease in X_{RG} results in a decrease for ΔK_E of \$1.7118742, until X_{RG} reaches \$462,000. Beyond this point the decrement in ΔK_E is changed to \$1.7124218 per dollar decrease in X_{RG}. From these data we construct a chart which shows the relationship between the sales limit and the net addition to retained earnings, or equivalently, the relationship between X_{RG} and ΔK_E when everything is adjusted in such a way that ΔK_E is maximum under the given conditions. This relationship between

X_{RG} and maximum ΔK_E is given by the upper line segments in the following diagram. The scale on the horizontal line has been adjusted to represent physical sales volumes rather than cost of sales.

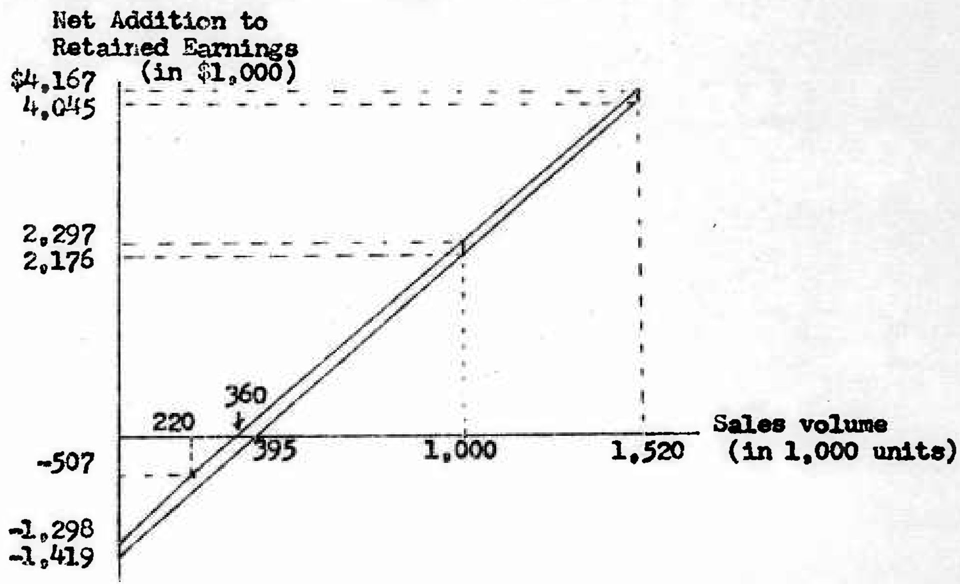


Figure VI-1

Relationship between Sales Volume and Net Addition to Retained Earnings

The lower line of Figure VI-1 was prepared as follows: Suppose that the management of this firm wants to know what the situation would be if their subordinates take the <u>worst</u> possible course of action they can do without violating the constraints imposed upon them. This means that management wants to know how much the subordinates can push ΔK_E down if they try to minimize it.

To derive this, we need to replace the minimum cash balance constraint (6.16) by a maximum cash balance constraint such as

(6.16a) $X_{CB} + X_{CR} + X_{CL} + X_{CE} - X_{BC} - X_{PC} - X_{DC} - X_{LC} \leq 8{,}000{,}000 - K_C$,

in which the maximum cash balance is constrained to be at most equal to $8,000,000. Then we replace the objective function (6.20) by its negative

(6.20a) Maximize $(-\Delta K_E =) - X_{CE} - X_{RE} - X_{GE} + X_{EF} + X_{EP} + X_{ED}$

which is equivalent to

(6,20b) Minimize $(\Delta K_E =) X_{CE} + X_{RE} + X_{GE} - X_{EF} - X_{EP} - X_{ED}$.

The minimum value of ΔK_E is - $1,419,050, which is derived by making no sales; maximum production for surplus inventory; no collection of accounts receivable; no payments on accounts payable incurring interest penalty, maximum amount of investment on government securities (which has the least interest rate among all methods of financing), and maximum amount of borrowing from a bank without violating the maximum cash balance constraint.

Now let us change the sales limit constraint (6.9) to

(6.9a) $X_{RG} \geq 0$

The dual evaluator associated with (6.9a) when the new problem in which (6.9) is replaced by (6.9a) is solved is - $1.7117638. This means that when the minimum sales limit in (6.9a) is increased by one dollar $-\Delta K_E$ is decreased or, equivalently, ΔK_E is increased by $1.7117638. This dual evaluator does not change for any value

of X_{RG} in the range of $0 \leq X_{RG} \leq 3,192,000$, and for $X_{RG} > 3,192,000$ the problem becomes unsolvable. From this information we derive the lower line in Figure VI-1.

Notice the significance of the information that this chart conveys. For example, suppose that management observes an actual sales volume of 1,000,000 units. Then, it can be sure that no matter how badly subordinates operate their divisions, the management is guaranteed that the net addition to retained earnings is at least as great as $2,176,000. Alternatively no matter how well the subordinates operate their divisions the net addition to retained earnings is at most $2,297,000. Of course, it is necessary to make sure that none of the constraints imposed upon subordinates are violated. But if this can be checked, e.g., by auditing and other internal control devices, the management can rely on the sales volume as a quick measure of the net addition to retained earnings.

On the other hand, if management is concerned with the breakeven point, this, too, can be translated into a range for the sales volume by means of this chart. In this case, the "breakeven sales range" is between 360,000 units and 395,000 units. This means that if the sales volume is less than 360,000 units, the management can be sure that the breakeven point has not been attained, whereas if the sales volume is greater than 395,000 units, it is sure that the breakeven point has been attained. Therefore, if it is management's goal to attain the breakeven point (which is a "satisficing" goal), it can be

translated into a sales volume which is given to his subordinates as their goal. He may use the minimum sales volume (360,000) or the maximum sales volume (395,000) or any figure in the range for the breakeven sales volume for this purpose, depending upon whether he is optimistic or pessimistic about the behavior of his subordinates.

Note that the process of deriving Figure VI-1 discussed in the above is only intended to describe the nature of the result in Figure VI-1, but it does not mean that we have to go through all these elaborate computations. By using the algorithm described in Appendix C, we can derive all needed information through a computer with minor modifications in a machine program for the simplex method.[1]

---

[1] We have developed the above analysis assuming that negative income taxes and negative dividends can be interpreted as deductions from the beinning balance of Dividends and Taxes Payable account, since the beginning balance of this account ($K_D = 2,922,000$) is large enough to cover negative dividends and taxes generated by even greatest possible deficit in the current period, i.e. the maximum deficit before taxes and dividends \$3,112,000 (\$1,419,050 / .456), (from which negative taxes and dividends which offset previous accruals) are calculated as \$1,693,000 (\$3,112,000 x .544). However, if this is not permissible, we can represent the asymmetric behavior of income taxes and dividends by means of methods described in Chapter II, Section 4 and in Chapter III. See also Charnes and Cooper, 1961.

---

2. A Sensitivity Analysis

We now move to a sensitivity analysis of the problem, by manipulating the selling price of the product as incorporated in the linear programming formulation through the constraint (6.5).

We shall analyze the relationship between selling price and sales volume when everything is adjusted optimally to attain the maximum net addition to retained earnings as follows. Let q be the sales volume in physical unit and p be the selling price of the product. To save space, we shall give only the result of our computation which shows the relationship between p and q when everything is adjusted optimally:

For $0 \leq q \leq 220{,}000$

(6.21) $$p = 11{,}998{,}636 \frac{1}{q} + 2.1$$

and for $220{,}000 \leq q \leq 1{,}520{,}000$

(6.22) $$p = 11{,}998{,}013 \frac{1}{q} + 2.10252$$

That is, a coefficient in a constraint (the coefficient 3.76 in (6.5) which was transformed into p) is now represented by a piecewise linear function of the inverse of the stipulation level (the stipulation in (6.5) transformed into sales volume q.)

For a practical range of p (e.g. $p \leq 20$ -- note that the current selling price is \$10 or, more precisely, \$9.996), the following chart was prepared from (6.22).

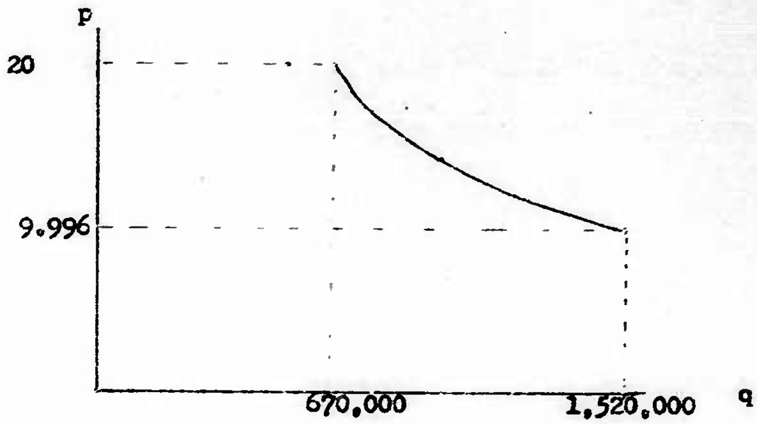


Figure VI-2

Selling Price and Sales Volume
at Optimum Solutions:
An Iso-Profit Curve

The significance of this chart may be seen as follows. Any point on this curve shows a selling price and sales volume pair which attains exactly the same amount of the net addition to retained earnings as given in Table VI-7, which is the maximum attainable amount when p = 9.996, assuming that everything is adjusted to attain maximum addition to retained earnings under the given constraints. This means that if the management can find a selling price at which they are sure that they can sell more than the sales volume indicated by the curve, the selling price has to be adjusted to that new price since this action will bring greater profit. That is, when a demand curve of the product is superimposed upon this iso-profit curve, any point in the demand curve which lies above the iso-profit curve shows a

preferable selling price than the current selling price of $9.996.

This kind of a sensitivity analysis can be extended in various directions, including a sensitivity analysis on cost factors which will help management in setting up standard costs. We obviously cannot shows all such possibilities of sensitivity analyses for managerial uses in this section. But, as the preceding discussion at least indicates, this kind of byproduct from our linear programming spread sheet model at least warrants further accounting research looking towards its practical use by management.

6. Multiple Period Model

Following the analyses of a single period model of spread sheet planning developed in the previous sections, we shall discuss some of the implications of extending the model to a multiple period model of spread sheet planning.

1. Goals

As we have discussed in the earlier chapters, we can incorporate a variety of goals, single or multiple, optimizing or satisficing, into our spread sheet planning model. For example, the goal may be to maximize the present value of the discounted future net income after taxes, or it may be to achieve the profit level at least as great as a target profit level in each period. In the former case we will need to specify the discount rate to be used, even though the effect of changing the discount rate on the solutions may be analyzed by means of a sensitivity analysis as discussed in the previous section. In the latter case we need not specify the discount rate, or we can use the consequences of a dual evaluator analysis, if we like, to study the discounting implied by any such target.

If satisfying multiple goals, (e.g., achieving a profit level at least as great as a specified minimum profit level in each period), are found to be incompatible, then we may proceed to an ordering of the goal of profit level achievement in each period. The management may be concerned with achieving a satisfactory profit level in the current period at the expense of unsatisfactory profit

levels in the future periods, or it may, in some cases, assume the reverse point of view.

Furthermore, a management may also respond as it sees some degree of substitutability between the goals of the satisfactory profit level achievement in each period in a dual evaluator analysis. Nor need it accept the dual evaluator values *per se*. The goals may, instead, be weighted according to what the management considers the relative degree of "regret" as we have discussed in Chapter III. Notice that in such a weighting we need not be concerned with the discount factor for the income stream, since if we like, an external market interest rate on investment may be incorporated in the model by means of a set of constraints (e.g., the constraint (6.2)). or in the functional, when the opportunities of outside borrowing and lending are to be considered. Therefore, in setting up the management goals for our model we need not be constrained by such unduly narrow concepts like maximizing the discounted future income stream[1] relative to only one rate of interest over all periods, though this may certainly be a valid goal in some situations.

1/ Actually recent studies have begun to reverse earlier econometric findings on the insignificance of interest rates as economic magnitudes only by introducing balance sheet considerations via such ideas as maintenance of an optimum portfolio of fixed asset relations. See, e.g., F. S. Hammer, 1963.

2. Subgoal

In order to extend our single period model to a multiple period model for n periods, we need to increase the number of spread sheet variables at least by a factor of n. It may need to be increased more if the nature of the constraints require finer divisions of the spread sheet variables.1/ On the other hand, it may be decreased

1/ There are, however, possible compensations via improved algorithms. See, e.g., Charnes and Cooper, 1961, Appendices F and G, or G. B. Dantzig and P. Wolfe, 1960.

by means of grouping the variables and simplifying constraints for remote future periods if the management is more concerned with the goals for the current and nearby future periods.

3. Constraints

In a multiple period model, we have two types of constraints, "intra-period" constraints and "inter-periods" constraints. The former are represented by constraints on operations in a given period, whereas the latter are represented by constraints connecting the operations in more than one period.

Let x^k be the vector for the spread sheet variables in the k^{th} period, and let

(6.23) $$B^k x^k \leq h^k \quad (k = 1, 2, \ldots, n)$$

be the "intra-period" constraints for the k^{th} period. Here we employ matrix notation so that we can thereby avoid a multiplicity of subscript. Then we can in this notation let

$$(6.24)\qquad \sum_{k=1}^{n} \hat{B}^k \hat{x}^k \leq \hat{h}$$

be the "inter-period" constraints over all periods.

In this approach we would have

$$(6.24a)\qquad \begin{bmatrix} B^1 & & & \\ & B^2 & & \\ & & \cdot\cdot & \\ & & & B^n \\ \hat{B}^1 & \hat{B}^2 & \cdot\cdot & \hat{B}^n \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ \cdot \\ \cdot \\ \cdot \\ x^n \end{bmatrix} \leq \begin{bmatrix} h^1 \\ h^2 \\ \cdot \\ \cdot \\ \cdot \\ h^n \\ \hat{h} \end{bmatrix}.$$

A particular kind of inter-period constraint relation is of interest to us because these arise as constraints on asset-equity balances, (e.g., (6.16) - (6.19)) in spread sheet balances. These may be represented as follows:

$$(6.25)\qquad \sum_{k=1}^{j} \hat{\hat{B}}^k x^k \leq \hat{\hat{h}}^j \qquad (j = 1, 2, \ldots, n).$$

If we now introduce these between (6.23) and (6.24) the matrix, with these submatrices as its elements, becomes

$$
(6.26) \quad \begin{bmatrix} B^1 & & & \\ & B^2 & & \\ & & \cdot\cdot & \\ & & & B^n \\ \hat{B}^1 & & & \\ \hat{B}^1 & \hat{B}^2 & & \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \\ \hat{B}^1 & \hat{B}^2 & \cdot\cdot & \hat{B}^n \\ \hat{B}^1 & \hat{B}^2 & \cdot\cdot & \hat{B}^n \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \\ \cdot\cdot \\ \cdot\cdot \\ x^n \end{bmatrix} \leq \begin{bmatrix} h^1 \\ h^2 \\ \cdot\cdot \\ h^n \\ \hat{h}^1 \\ \hat{h}^2 \\ \cdot\cdot \\ \hat{h}^n \\ \hat{h} \end{bmatrix}
$$

4. <u>Solutions</u>

Solutions from the multiple period model may be used in a variety of ways, some of which have already been discussed in Section 3. They may be used to prepare projected balance sheets and income statements for each period under consideration. Also the solutions obtained for the components of any x^j may be given to the subordinates as targets in particular periods. In the latter case, we may discard the values of the spread sheet variables for the future periods derived by a solution and focus on the spread sheet variables only for the current period. Then, at the end of the current period, we review the constraints and goals, modify the formulation, solve the new problem, and then use the values thus determined for the spread sheet variables only for the current period at that time.[1]

[1] This, to be sure, will generally be only a sub-optimization in the sense of C.J. Hitch (1953) -- but the dual evaluator analysis can also play a significant role here in reducing the degree of sub-optimization.

5. Analyses

The dual evaluator analysis and its extension as well as the sensitivity analysis elaborated in the previous two sections can, of course, also be applied to the multiple period model, too. Furthermore, the concept of "transfer pricing" can also be applied to the multiple period model, since the balances of asset-equity accounts at the end of the k^{th} period are mathematically the same as products transferred from the k^{th} division to the $(k+1)^{st}$ division in the linear programming formulation.[1/] Thus, accounting

---

1/ Refer to Whinston, 1962, for the discussion of transfer pricing.

---

theory of asset valuation based on opportunity costs can readily be developed in this direction, and, moreover, this can be done by (a) introducing ordinary accounting costs, considered as projected outlays, into the direct model, while (b) securing the corresponding opportunity costs from the results of formulation and solving a corresponding linear programming model. On the other hand, this does not mean that all aspects of current procedures should be observed. The usual distinctions between capital and operating budgets are a case in point. This distinction may best be viewed as one which arises because of the limited ability of current accounting procedures when confronted with large and complex problems involving many interacting variables.

It is not something which is required by the logic of double entry accounting per se.

6. Capital budgeting

While we cannot exhaust here all possibilities of the analyses of a multiple period model for spread sheet planning, we shall discuss briefly one further important direction in terms of possible applications.

Let us consider a set of plans for the expansion of production capacities. We assume that the production capacity constraints have already been incorporated in the model as "intra-period" constraints, e.g., (6.7). We, then, decompose the related fixed asset accounts (and consequently spread sheet variables, too) to provide enough detail so that the dollar value of the assets can be related with sufficient closeness to the production capacity alterations being considered. Once we have prepared a decomposed spread sheet variable which represents the relationship between a dollar value increment and a production capacity increment, we can incorporate the variable into the production capacity constraints with an appropriate coefficient for the conversion of a dollar value increment into a production capacity increment.

For example, in our single period model developed in Section 2, suppose that the fixed assets consist of a set of homogeneous machines, and that an addition of one unit of machine which costs \$100,000 increases the monthly production capacity by 5,000 units, or \$5,000 of raw materials at standard cost. If we represent any amount of machine purchases in the k^{th} period by X_{FP}^{k}, then the production capacity constraint (6.7) for the j^{th} period may be changed into

$$(6.7a) \qquad X_{GM}^{j} - .05 \sum_{k=1}^{j} X_{FP}^{k} \leq 1,300,000. \qquad (j = 1, 2, \ldots, n)$$

If we want to impose a time lag between the period of purchase and the period in which the production capacity increase is effected, we may represent this by the following constraint

$$(6.7b) \qquad X_{GM}^{j} - .05 \sum_{k=1}^{j-s} X_{FP}^{k} \leq 1,300,000 \qquad (j = 1, 2, \ldots, n)$$

where s is the number of periods of time lag and $X_{FP}^{k} \equiv 0$ for $k \leq 0$. In an entirely analogous way, a decrease in production capacity due to the disposal of the machines may also be effected.

Of course, in many cases, the purchases of a fractional unit of machine does not have any meaning, but this may be adjusted by applying techniques similar to the one used in dealing with fixed costs in Chapter II or by so-called diophantine (integer) linear

programming techniques.[1/] Also, if the relationship between the

1/ C.f. Charnes and Cooper, 1961. Also see Weingartner, 1961.

dollar increment of the fixed assets and the production capacity increment is not linear, the method described in Chapter II for dealing with a piecewise linear cost curve can be applied here. In addition, if we are interested in a particular investment plan which will result in an expansion of the production capacity or in cost savings, the dual evaluator analysis and its extension, as well as the sensitivity analysis (on cost factors), which have been discussed in the previous sections should supply valuable information for the decision making about the investment.

In summary, a multiple period model of spread sheet planning carries with it the implication that capital budgeting and operating budgeting are to be considered <u>simultaneously</u>. That is, the model derives the solutions which, among various alternatives for capital budgeting and operating budgeting, give us the minimum degree of "regret" in the management goals. It therefore instructs us to invest only after taking into account such factors as depreciation,[1]

[1]Or, rather, its funds-flow consequences.

income taxes, investment alternatives, sales constraints, as well as management goals and management utility (degree of regret) in the attainment of the goals, not all of which are considered in an ordinary capital budgeting procedure. Thus, by means of such

a multiple period model as described above, the usual artificial separation of the two types of budgets, capital budget and operating budget, is avoided and the funds flow consequences can then be considered <u>across</u> as well as <u>within</u> these budgets.

CHAPTER VII

SUMMARY AND CONCLUSION

1. Summary of the Results of the Studies

In this concluding chapter, we shall first briefly summarize the results obtained from our studies and then, in the next section, we shall attempt some assessment of existing accounting practices -- at least as we have experienced them or read about them -- in order to determine how they are functioning as "management information systems" and then, finally, to suggest possible ways in which they might be improved for the latter purpose.

Using numbers in parentheses to refer to the chapter, section, and subsection, (in this order), in which a result is discussed, we summarize the main results of our studies as follows:

1. A planning process may be characterized as a translation of a set of given goals into a set of subgoals which are more operational and controllable. (I-2, II-1-(1)).

2. Goals may be "optimizing" or "satisficing," and single or multiple. (II-1-(1)).

3. Breakeven analysis may be viewed, essentially, as a way to translate a goal (profit level) into a more operational subgoal (sales volume). It carries with it the notion of satisficing (no-profit no-loss point). (II-1-(1)).

4. The breakeven model can be extended to a piecewise linear breakeven model which takes into account nonlinear revenue and cost curves. A nonlinear cost curve may contain such factors as fixed, variable, semi-fixed, and semi-variable costs, but these can all be represented by a set of "variable cost components." (II-1-(2)).

5. Such piecewise linear representations are possibly more closely akin to accounting rather than economics models and modes of analysis. But this does not mean that basic ideas in economics which are potentially useful need to be abandoned. This is illustrated by the idea of elasticity of demand which has been developed in economics. Piecewise linear models cause difficulty with the ordinary constructs which are not defined at kinks in a piecewise linear total revenue curve. But this, too, may be viewed only as a methodological difference. The basic idea may be preserved by suitable means. This leads to ideas like our concept of "marginal revenue head." (II-1-(3)).

6. Models for a planning process which involves translation of a single goal into multiple subgoals, may be treated by means of generalized inverses and goal programming. (II-3, II-4). This extends to constraints imposed on subgoals as well as the constraints which operate on the goals. The two may be simultaneously treated and contradictions, if any, resolved by various minimization measures.

7. Varying the functional in a goal programming model produces various types of goal orientation, satisficing, optimizing, etc. (III-2).

8. Models for multiple as well as single goals may be treated in these same kinds of functional representations. (III-3-2, III-4-1).

9. When multiple goals are incompatible, the goal programming model and generalized inverse models both minimize the deviation from goal levels. In the goal programming model the deviation is defined as the sum of the absolute value deviations for each goal. In the generalized inverse model the deviation is defined as "Euclidean distance" -- minimum sum of squares -- from the goal levels. ((III-4-(2)).

10. Treatment of goal incompatibilities is not restricted to a scalar treatment and commensurability of different goals. Models may be constructed which incorporate an "ordering" of multiple goals without raising issues of goal commensurability. Alternatively a "weighting" of multiple goals may be used as when the commensurability of goals is derived by means of the degree of management's relative "regret" on one unit of unsatisfactory deviation from the goal level. (III-3-(3), III-4-(3)). Such a relative "regret" may also depend upon the actual distance from the goal level, in which case a step function may be incorporated in the goal programming model. (III-3-(3)).

11. Accounting processes are assessed and studied relative to the subject of "indicators" on goals and subgoals at both planning and control stages. These indicators are related to elements in a vector w, a matrix C, and to a vector x for subgoal variables. (IV-2).

12. If the goal variables represented by a vector v is related to the subgoal vector x by $Ax = v$, with A a matrix conformable to x, and if the indicator vector w is related to this same subgoal vector x by $Cx = w$, the necessary and sufficient condition for v to be a uniquely determined function of w only is that each row of A is expressible as a linear combination of the rows of C. If an indicator or a set of indicators satisfies this condition it is called a "perfect" indicator or a "perfect" indicator set. (IV-3).

13. "Imperfect indicators" are the class of all indicators that are not perfect. This is studied relative to possible constraints that may be imposed in order to narrow the range of goal indicator discrepancy. The maximum and the minimum value of the goal level at each indicator value can be represented by an "indicator-goal control chart." (IV-4).

14. Such a chart can be produced by a mapping of the original convex set (generated by a set of constraints on the subgoal variables) over the two dimensional space of indicator values and goal levels. An algorithm to perform such a mapping and generate all the extreme points in the mapped convex set, from which an "indicator-goal control

chart" can be prepared, is developed in Appendix C. The chart may also be prepared as an extension of a dual evaluator analysis. (VI-5-(1)).

15. This indicator-goal control chart can tell management (i) at any value of the indicator how much they have attained, in terms of the goal, at best and at worst, (ii) for any given goal level what value of the indicator guarantees that the goal level has been attained and what value of the indicator guarantees that the goal level has not been attained, and (iii) if a subordinate's goal is taken as an indicator, whether, and by how much, management goals and subordinate's goals may be in conflict. (IV-4, VI-5-(1)).

16. Choice of an indicator for any given goal depends upon the goal function and the constraints imposed upon the subgoal variables. In order to aid in dealing with this kind of problem an "indicator-goal divergence coefficient," or simply a "divergence coefficient," is developed to show the degree of divergence between a goal and any indicator that may be used. (IV-5).

17. The basic equations of double-entry bookkeeping can be restated and developed in order to obtain suitable mathematical aids for applying these goal analysis models and ideas by reference to the data and procedures utilized in ordinary accounting practice. (V-2). We emphasized, in particular, the fact that multiple measuring units can perfectly be used in the double entry bookkeeping system. (V-2).

These are further extended to secure added power and flexibility for planning as well as control aspects of accounting by studying relationships between transactions and asset-equity balances. The accounting idea of a spread sheet is brought together with the mathematical idea of an incidence matrix. The latter has known relationships to the idea of networks. Hence by this means the idea of an accounting network is rigorously established in a way that provides convenient access to the related graphical representations. (V-3).

18. Taking elements in a spread sheet as subgoal variables, which are called spread sheet variables, the networks are related to equivalent linear programming models for analyses of goals. These are then applied to accounting data by means of specific examples which illustrate their possible use for ordinary accounting functions. (V-4).

19. Planning by spread sheet variables, or "spread sheet planning," is developed around the topic of planning on aggregates of subgoal variables in which spread sheet variables are used for aggregates. Therefore, the appropriateness of such a planning on aggregates depends upon the relationship between the goal function and the aggregation process in the accounting system. These topics are discussed and then by reference to cases such factors as the divergence coefficient, introduced earlier, are brought into the picture. (V-5).

20. An experimental application of the idea of "spread sheet planning" was effected by means of a firm's operations in the Carnegie Tech Management Game. This sufficiently complicated game allowed us to incorporate various phases of business operations, such as production, marketing, finance, etc., into our model of spread sheet planning (VI-2).

21. For the latter analysis a single management goal was translated into values for a set of spread sheet variables. The latter are available, then, as goals which may be given to subordinates. Alternatively the spread sheet variables may be combined to obtain a projected balance sheet and a projected income statement as important documents often synthesized by accountants as an aid to overall management planning. (VI-3).

22. These balance sheet and profit and loss statements may be viewed as projections in terms of outcome possibilities in ordinary accounting treatments of the corresponding transactions. These are sometimes contrasted with alternative opportunity cost treatments. However, here the latter are viewed as being available from the same model differently interpreted. To illustrate this an analysis was made on a dual evaluator, which was summarized in the form of an accounting statement for opportunity costs. (VI-4). The dual evaluator analysis was extended further by means of an indicator-goal control chart. (VI-5-(1)). A sensitivity analysis was made on the selling price of the product and the result was represented by an (optimum) iso-profit curve

which shows the relationship between selling price and sales volume at a given profit level when everything is adjusted to attain maximum profit. (VI-5-(2)). The purpose of these analyses is to show how managerially valuable information can be derived from accounting data when they are combined with mathematical models such as those developed in earlier chapters. Also the fact that these analyses are not of the "ceteris peribus" variety but of the "mutatis mutandis" ones was emphasized. (IV-4, IV-5).

23. Implications of extending the above single period model into a multiple period model were discussed. It was noted that in a multiple period model, a discount rate for income stream, etc., need not be specified, since satisficing multiple goals, e.g., attaining a profit level at least as great as the minimum profit level in each period, can very well be incorporated in the model. In this case, if the multiple goals are found to be incompatible, they may be ordered or weighted as discussed in Chapter III. In addition, weighting of unsatisfactory profit levels in each period may be made by reference to the degree of management psychological relative "regret," and the resulting interest rate imputation can be compared, if desired, with market rates of interest. (VI-6-(1)).

24. The multiple period model for spread sheet planning was also examined to cast further light on accounting ideas which are basic against merely procedural approaches which, even if widespread,

may only be a result of limitations imposed on the methodology available for implementing the more basic constraints. This was illustrated by reference to the possibilities of joining capital budgeting simultaneously with operating budgets as an alternative for the two types of budgets. (VI-6-(6)).

2. Conclusion

Taking into account these results we now review some aspects of current accounting practice from the standpoint of "management information systems." This is coupled with some statements on how accounting might now be developed with the newer methodologies and ideas that are now becoming available.

1. Accounting for Control

Accounting is ordinarily classified into financial accounting and management accounting. This is convenient in some contexts. For our purposes, however, we found it better, in Chapter I, to distinguish between "accounting for planning" and "accounting for control." This enabled us to isolate certain criteria in an accounting system design and also to better understand some aspects of current practice. Thus in accounting for planning it was found that flexibility of records provided a criterion in order that a wide variety of possibly non-repetitive demands for data might be readily supplied with minimum time or costs. In accounting for control the criteria of uniformity and consistency come to the fore. Thus, from the standpoint of external control of management itself, even financial accounting may be better understood in terms of its heavy reliance on these criteria. But this does not end here and, in fact, the bulk of the control problems we discussed lie in (internal) management accounting. This, too, provides the focus of our following discussions.

As mentioned in Chapter I, it is rather difficult to construct systematic theories of accounting for planning due to non-repetitive and even non-accounting (double entry) varieties of the data that may be required by management for planning purposes. On the other hand, something may be done here, too, if we approach this topic from the standpoint of its possible relations to control accounting. The latter is evidently more easily treated since it has uniformity and consistency as bases to work from. It also has relatively well developed concepts such as standards, measurement units, scales, meters, etc., all requiring uniformity and consistency. Working therefore from this point of view we have tried to supply a basis for a more general approach to control accounting theory as well as a basis for implementation. However, more extended constructions of systematic theories can be and should be made in accounting for control.

Of course, there are also other functions such as "stewardship" accounting, as well as accounting for taxes, and so on. But these are not of major interest here and so we have focused our attention and will focus our attention in the following discussion only to the function of control accounting that has been discussed in this thesis.

2. <u>Coordination between Planning and Accounting for Control</u>

In Chapter IV, we analyzed the relationship between planning processes and control accounting processes and derived a concept

of "divergence" between the two processes represented by an "indicator-goal divergence coefficient." It is a truism in accounting that knowledge of management goals is essential. Sometimes, however, this topic is confused with others and, in particular, certain kinds of accounting methodologies may be allowed to assume a precedence in this respect. Hence it is desirable to proceed at a more fundamental level to isolate these aspects of accounting from accounting ideas that are really basic. This thesis has been conceived therefore with this as a topic for theoretical clarification as well as for extending the power of the basic accounting ideas.

Of course this thesis is only a start in this direction. We have, however, carried it far enough, we think, to indicate that it should be possible to develop adequate accounting methods to meet all kinds of management goals without in any way altering the really basic principles of accounting. This is in contrast, we might note, with some of the recent efforts to secure a better basis of servicing management by stepping outside the conventional realm of accounting and by exploring the nature of business operations and internal and external environment in which these operations take place.1/

1/ See, e.g., Moonitz, 1961; Sprouse and Moonitz, 1962; Chamber, 1961, etc.

It is, of course, only trite to say that accountants should carefully study the requirements of management and then design their accounts and reports relative to the way individual activities may be identified with related goal attainments. It is not trite, however, to attempt to delineate ways in which this can be done relative to the evaluation and aggregation of subgoal activities according to their contributions to the attainment of the goals. It is this latter topic, of course, that has been a major objective for this thesis. We have tried to do this moreover in a way that will open paths to future progress in accounting as well as in management. This, too, is not trite if one views this as an issue in developing wholly effective ways of achieving this service to management while retaining the basic aspects of accounting practice which have now served so well over a period of centuries.1/

1/ Refer, for example, to the following quotations:

"Capitalism without double-entry bookkeeping is simply inconceivable. They hold together as form and matter. And one may indeed doubt whether capitalism has procured in double-entry bookkeeping a tool which activates its forces, or whether double-entry bookkeeping has first given rise to capitalism out of its own "rational and systematic spirit." (Werner Sombart, 1922).

"That the countries in which the science of bookkeeping made the most progress were always those in which most economic progress was being made can no doubt be explained as a mixture of cause and effect." (H. M. Robertson, 1935).

3. Control Accounting for Non-profit Goals

From the beginning of our study we have not limited our analyses so that they will be applicable only to particular profit management goals -- e.g., maximum profit or satisfactory profit. We have also not even restricted ourselves only to particular goal and subgoal structures. We have instead attempted to delineate a general route for dealing with all kinds of possibilities here in such a way that neither down-the-line nor top management goals would have to be ignored. For instance, we have considered it to be perfectly natural for a sales manager to try to reach a goal level of sales volume even if the top management has only overall profit for its goal. If it is the sales manager's goal to attain a given level of sales volume, then accountants must provide him an appropriate feedback which tells him, from time to time, where he stands in terms of this specified goal. On the other hand we have attempted also to deal with these ideas in a way that can help such a sales manager, as well as top management, to understand the implications of goal alterations. The dual evaluator analysis provided an illustration. Furthermore, we have not attempted to do this piece by piece, or <u>ceteris paribus</u>, but rather we have dealt with these ideas relative to contexts in which a host of possible intergoal interactions must be considered <u>mutatis mutandis</u>. Here, too, it is our belief that much more can and should be done by further developments in these directions.

There are further alternatives, too, relative to current controversies on the monetary bases of accounting. It is partly with this thought in mind that we introduced a matrix reformulation of the basic accounting equations. We did not exploit this here, however, for its bearing on monetary valuation. Hence we can now only note that the convention of monetary valuation may itself only be treated in full perspective after further study of other possibilities like these. From a control standpoint in any event, this one currently pervasive idea of monetary valuation may better be replaced by a more general theory of valuation and aggregation, and here the "divergence coefficient," may find a place as we turn to a variety of goals and subgoals.

There are further issues of time-period consequences for goals and subgoals that require more attention than we could accord them. Instead of restricting our research on accounting to topics like asset-equity valuations only -- and there has been ample discussion of this in the literature of accounting -- we should also begin to study <u>activities</u> that attain a goal so that we can properly begin to give requisite attention to activities rather than assets and equities which represent the <u>results</u> of activities.[1/] This does

---

1/ Some of the information alteration schemes suggested by A. Stedry, 1960, and by A. Stedry and A. Charnes, 1962, are pertinent here and so are some of the procedural studies studied by N. C. Churchill, 1962, and N. C. Churchill and W. W. Cooper, 1962.

not mean that assets and equities should be eliminated from accounting records. It only means that we can best gain perspective on fundamental principles which govern the recording of assets and equities by studying them in a broader context of activity accounting.

4. Accounting and Mathematics

A good deal more can be said on these topics. But now we want to turn to a final discussion of the role that mathematics might play in developing theories of control accounting in a higher dimension.

Throughout this thesis we have employed mathematics not only, hopefully, as a procedural device for developing better accounting for management but also as a device for securing theoretical clarification. Finally, we have also used it to develop firm contacts not only with mathematics (e.g., linear programming) but also with other disciplines as well (e.g., network theory). It is our belief that mathematics was indispensable in achieving these kinds of clarifications and procedural extensions. But it cannot be denied, in any event, that mathematics was indispensable in order to establish wholly rigorous -- hence firm -- contact with these other disciplines. Here, too, much more can and should be done. Notice, for instance, that rigorous connections have been established between linear programming and statistics and the theory of games. This is a

direction of further research which will be needed if we are going to develop satisfactory accounting approaches to topics like goal control under risk and uncertainty.

If we look back on the history of accounting, we can note that the development of accounting was rather closely attended to by the work of other mathematicians, and, indeed, the first publication of double entry accounting appeared in a tract on mathematics.[1/]

1/ See Charnes, Cooper, and Ijiri, 1962.

The advent of the electronic computer, a mathematics tool, and the development of new mathematics along with numerous management applications in operations research and elsewhere open further prospects for future developments. It is our belief that serious research will show that accounting is neither in conflict with these developments nor isolated from them. As a case in point we can observe the recent development and use of network ideas for scheduling and controlling projects. These, too, can be rendered in accounting format and shown to have the basic double entry principle embedded in them. The fact that these methods of approach do not fit easily into the "T account" arrangement which is customary in accounting is not a matter of fundamental importance. The mathematics we have been using lends itself readily to these new usages[2/] as accounting

---

2/ See, e.g., A. Charnes and W. W. Cooper, 1961, for a precise characterization of so-called critical path scheduling as a

(Footnote 2 continued)

special kind of linear programming model which is intimately related to the kinds of accounting networks that we examined in Chapter V. See also Charnes, Cooper, and Thompson, July, 1962, for a subsequent extension of critical-path and PERT techniques into the realm of statistical (risk) considerations.

---

possibilities and it also lends itself [illegible] spread sheet recording and reporting on electronic computers.

5. Accounting in the Future

Planning processes and control accounting processes will be better coordinated, and the function of accounting as a "management information system" will best be improved, we think, only after preceding studies and careful reviews of accounting principles, theories, and practices, etc., at a sufficiently fundamental level. Even though control accounting is only one branch of accounting it is certainly a very important one. It also provides a good point of departure for further insight and development. There is a need and an opportunity, in any event, for further development of a rigorous accounting theory as well as better methods of servicing management. We have tried to show how both of these objectives can be accommodated. But, of course, this thesis is only a start toward all of the needs that must be met.

APPENDIX A

AN ANALYSIS OF GENERALIZED INVERSES

1. Introduction

A generalized inverse of a matrix A (denoted by $A^{\dagger}$) may be defined as the unique matrix which satisfies all the following four set of equations:

(A1.a) $$AA^{\dagger}A = A$$

(A1.b) $$A^{\dagger}AA^{\dagger} = A^{\dagger}$$

(A1.c) $$AA^{\dagger} = (AA^{\dagger})^*$$

(A1.d) $$A^{\dagger}A = (A^{\dagger}A)^*$$

where $(AA^{\dagger})^*$ and $(A^{\dagger}A)^*$ represent the transposes of the matrices $(AA^{\dagger})$ and $(A^{\dagger}A)$, respectively.[1] Penrose (1955) proved that for

[1]More generally, when the matrix A is complex, $(AA^{\dagger})^*$ and $(A^{\dagger}A)^*$ represent the conjugate transposes of the matrices $(AA^{\dagger})$ and $(A^{\dagger}A)$, respectively. However, since only real matrices are discussed in this appendix, conjugate transposes are reduced to ordinary transposes, and hence we shall be concerned with the latter only in our discussion.

any matrix A (non-singular or singular, square or rectangular, zero or non-zero) there exists the unique matrix $A^{\dagger}$ which satisfies above set of equations.[1]

[1]Alternatively, a generalized inverse may be defined as the unique solution X to $AX = P_{R(A)}$ and $XA = P_{R(X)}$ where $P_{R(A)}$ represents the orthogonal projection on the range of A and $P_{R(X)}$ the orthogonal projection on the range of X (Ben-Israel and Charnes, 1963)

Since any orthogonal projection is hermitian idempotent (i.e., $P = P^* = P^2$, where P^* is the conjugate transpose of P), this definition is equivalent to the definition given by using (A1). (See also Appendix B.

When a matrix A has its ordinary inverse A^{-1} (i.e., when A is non-singular), $A^\dagger$ is equivalent to A^{-1} since A^{-1} satisfies (A1) and the uniqueness property of $A^\dagger$ guarantees that $A^\dagger \equiv A^{-1}$.

In order to make the properties of a generalized inverse more visible, we shall work on particular examples. Let us take 2 by 2 matrices $\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}$ which will be denoted by B and C, respectively. Then, the transposes of B and C are $\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$, respectively. The matrix C has its ordinary inverse C^{-1} which is $\begin{bmatrix} 1.5 & -2 \\ -.5 & 1 \end{bmatrix}$, whereas the matrix B does not have the ordinary inverse since it is singular. The generalized inverse $B^\dagger$ exists, however, for the matrix B as for any other arbitrary matrices and is given by $\begin{bmatrix} .08 & .04 \\ .16 & .08 \end{bmatrix}$ as we shall see later.

In summary, we shall use the following matrices in our discussions.

	A singular matrix B	A non-singular matrix C
The original matrix	$B = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$	$C = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}$
The transposed matrix	$B^* = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$	$C^* = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$
The generalized inverse	$B^\dagger = \begin{bmatrix} .08 & .04 \\ .16 & .08 \end{bmatrix}$	$C^\dagger = \begin{bmatrix} 1.5 & -2 \\ -.5 & 1 \end{bmatrix}$
The ordinary inverse	(does not exist)	$C^{-1} = C^\dagger$

Notice that $B^\dagger$ and $C^\dagger$ satisfy (A1). The discussions in this appendix may be generalized to rectangular matrices since any $m \times n$ rectangular matrix may be converted into a (singular) square matrix by adding $(m - n)$ zero columns (if $m > n$) or $(n - m)$ zero rows (if $m < n$).

2. Linear Transformations

An m x n matrix A is uniquely associated with a linear transformation which assigns to any vector x in an n-dimensional vector space $\mathcal{X}_n$ the vector y = Ax in some subspace of an m-dimensional vector space $\mathcal{Y}_m$, satisfying the condition that

$$A(cx_1 + cx_2) = cAx_1 + cAx_2 \tag{A2}$$

for every x_1 and x_2 in $\mathcal{X}_n$ and every scalar c. The vector space $\mathcal{X}_n$ is called the domain of A and the subspace of the vector space $\mathcal{Y}_m$ the range of A.[1] For example, the matrix B represents a

[1]Refer, for example, to Kemeny, Mirkil, Snell, and Thompson, 1959, p.255 ff. and p.278ff.

linear transformation which assigns to any vector x (e.g. $\binom{3}{2}$) in a two-dimensional vector space $\mathcal{X}_2$ the vector y = Bx (e.g. $\binom{14}{7}$) in an another two dimensional vector space $\mathcal{Y}_2$. Similarly, the matrix C represents another linear transformation which assigns to any vector x (e.g. $\binom{3}{2}$) in $\mathcal{X}_2$ the vector y = Cx (e.g. $\binom{14}{9}$) in $\mathcal{Y}_2$.[1]

[1]Unless the distinction becomes important, we shall, for simplicity, use the symbols A, B, and C for the matrixes A, B, and C, respectively, and for the transformations represented by them. We shall also use the term points and vectors interchangeably.

Notice that the vectors in $\mathcal{Y}_2$ assigned by the transformation B are always in the form of $\binom{2\alpha}{\alpha}$ where α is a scalar no matter what vector we choose in $\mathcal{X}_2$. This means that no vectors in $\mathcal{X}_2$ correspond, under the transformation B, to the vectors in $\mathcal{Y}_2$

which are not in the form of $\binom{2a}{a}$. Furthermore, there are infinitely many vectors in $\mathscr{X}_2$ which correspond to the same vector $\bar{y}$ in $\mathscr{Y}_2$ under the transformation B. For example, let $\bar{x}$ be one of the vectors in $\mathscr{X}_2$ which corresponds to $\bar{y}$ in $\mathscr{Y}_2$. Then, any vector x given by

(A3) $$x = \bar{x} + \binom{-2}{1}z$$

where z is a scalar corresponds to $\bar{y}$ under the transformation B, since, for any choice of z,

(A4) $$Bx = B\bar{x} + B\binom{-2}{1}z = B\bar{x} = \bar{y}.$$

These properties of the transformation B cannot be observed in the transformation C. For any vector $\bar{y}$ in $\mathscr{Y}_2$, we can always find a vector $\bar{x}$ in $\mathscr{X}_2$ which corresponds to it, since $\bar{x} = C^{-1}\bar{y}$ corresponds to $\bar{y}$ under the transformation C ($C\bar{x} = CC^{-1}\bar{y} = \bar{y}$.) Furthermore such a vector $\bar{x}$ is unique for any given $\bar{y}$, since any vectors in $\mathscr{X}_2$ other than $\bar{x}$ which may be written as $\bar{x} + x$ $(x \neq 0)$ correspond, under the transformation C, to a vector in $\mathscr{Y}_2$ given by $C\bar{x} + Cx = y + Cx \neq y$ for $Cx = 0$ if and only if $x = 0$ which comes from the fact that C is non-singular.

In general, for the transformation given by an m x n matrix A, a set of all vectors Ax $(x \in \mathscr{X}_n)$ constitutes a subspace of $\mathscr{Y}_m$ which is called the range of A. We shall denote the range of A by R(A). The dimension of R(A), or the number of linearly independent vectors which span R(A), is called the rank of A which will play a crucial role in our discussions.

Then, for a vector y in $\mathcal{Y}_m$ but not in R(A), there is no vector in $\mathcal{X}_n$ which corresponds to y under the transformation A. This can happen if and only if the rank of A, denoted by r, is less than m. Furthermore if the rank of A is less than n, all the vectors in a subspace of $\mathcal{X}_n$ whose dimension is equal to n-r, correspond to the same given vector in the range of A. Therefore, for every vector y in $\mathcal{Y}_m$ the existence and uniqueness of a vector in $\mathcal{X}_n$ which corresponds to y are guaranteed if and only if m = n = r, i.e., if and only if the matrix is non-singular. These points will be elaborated later in more detail.

Let us look at these issues from a view point of <u>solutions</u> to a system of simultaneous linear equations. A system of simultaneous linear equations may be represented by $Ax = \bar{y}$ where A is an m x n matrix, x an n-dimensional vector of variables and $\bar{y}$ an m-dimensional vector of constants. Therefore, by solving a system of simultaneous linear equations we mean to find a set of vectors in $\mathcal{X}_n$ which correspond to the given vector $\bar{y}$ in $\mathcal{Y}_m$ under the transformation A. Then, the above arguments on the singularity and non-singularity of matrices can also be applied to the problem of the existence and the uniqueness of a solution to a system of simultaneous linear equations. I.e., the existence <u>and</u> uniqueness of a solution to a system of simultaneous linear equations are guaranteed for every m-dimensional vector y if and only if

the matrix associated with the system is non-singular. If the rank, r, of an $m \times n$ matrix A is less than m, there exist m-dimensional vectors, y's, for which $Ax = y$ fails to have a solution. Moreover, if the rank of A is less than n, the solution to $Ax = y$, if exists, is not unique.

3. One-to-One Correspondence Between R(A) and R(A*)

We have observed that the range of A, denoted by R(A), is a subspace of $\mathcal{Y}_m$ whose dimension is r, the rank of A. This means that any vector y in R(A) can be expressed by a linear combination of r basis vectors, i.e., r linearly independent vectors which span R(A). For example, the matrix B given in Section 1 has rank 1. R(B) is therefore a one-dimensional subspace of a two-dimensional space, $\mathcal{X}_2$. Hence, any vector in R(B) can be represented as a scalar multiple of one basis vector such as $\binom{2}{1}$. On the other hand, such a set of r basis vectors, or simply a basis, is not unique. (E.g. $\binom{1}{.5}$, $\binom{4}{2}$, etc., can be a basis, of R(B).) However, let us arbitrarily choose a set of r basis vectors and prepare an n x r matrix U. The columns of U provides a basis. (E.g., $U = \binom{2}{1}$.) Then, any vector y in R(A) can be represented as y = Uw where w is an r-dimensional vector. Moreover, by virtue of the fact that the r columns of U are a basis, there is a one-to-one correspondence -- a so-called isomophism -- between y and w, i.e., for any y in R(A) there is one and only one w such that y = Uw.

Similarly consider an n x m arbitrary matrix $\hat{A}$ whose rank is equal to r, the rank of A, and whose range, R($\hat{A}$), has r basis vectors, none of which are transformed into the null vector in $\mathcal{Y}_m$ under the transformation A. The latter condition is needed

to insure some properties necessary for $\hat{A}$ to be an "inverse" transformation of A as we shall see later.

We shall derive the generalize inverse of A from $\hat{A}$ by imposing some conditions upon A successively. First, note that any vector x in $R(\hat{A})$ can be represented as $x = V\hat{w}$ where V is an n x r matrix of bsis vectors for R(A) and $\hat{w}$ is an r-dimensional column vector for a linear combination of basis vectors. Again, by virtue of a basis, there is one-to-one correspondence between x and $\hat{w}$, i.e., for any x in $R(\hat{A})$ there is one and only one $\hat{w}$ such that $x = V\hat{w}$. Since w and $\hat{w}$ can be any vector in an r-dimensional Euclidean space, $\mathcal{Z}_r$, any r-dimensional non-singular matrix can give a one-to-one correspondence between a set of w vectors and a set of $\hat{w}$ vectors. This implies that if we restrict our attention only to vectors in R(A) and $R(\hat{A})$, we can make a one-to-one correspondence between vectors in R(A) and vectors in $R(\hat{A})$.

Now, let us consider a vector x in $R(\hat{A})$. Since $R(\hat{A})$ is a subspace of $\mathcal{X}_n$ we can apply the transformation A to the vector x obtaining the vector $y = Ax$ in R(A). We would like to find a transformation A which transforms y precisely back to x which we started with. What condition

should we impose upon the matrix $\hat{A}$ given the original matrix A? Let us take an arbitrary x in $R(\hat{A})$. We transform x into $y = Ax$ by the transformation A and then transform y back to a vector in $R(\hat{A})$ by the transformation $\hat{A}$. The vector thus obtained is $\hat{A}Ax$. We want to have this equal to x. Therefore, one of the conditions is

(A5) $\qquad \hat{A}Ax = x \quad \text{for all } x \in R(\hat{A})$.

Similarly, starting with an arbitrary vector y in R(A) we apply the transformations $\hat{A}$ and A successively and require the result to be exactly equal to y, i.e.

(A6) $\qquad A\hat{A}y = y \quad \text{for all } y \in R(A)$

Since $R(\hat{A}) \equiv \{Ay: \; y \in \mathcal{Y}_m\}$ and $R(A) \equiv \{Ax: \; x \in \mathcal{X}_n\}$

(A5) and (A6) may be written as

(A7) $\qquad \hat{A}A\hat{A}y = \hat{A}y \qquad y \in \mathcal{Y}_m$

(A8) $\qquad A\hat{A}Ax = Ax \qquad x \in \mathcal{X}_n$

or

(A9) $\qquad \hat{A}A\hat{A} = \hat{A}$

(A10) $\qquad A\hat{A}A = A$.

It will be clear from (A1.a) and (A1.b) that the generalized inverse of A satisfies both (A9) and (A10).

A solution $\hat{A}$ to (A9) and (A10) does exist for any arbitrary matrix A as proved in Penrose, 1955; however it is, in general, not a unique solution since we ignored two other constraints on A as given in (A1.c) and (A1.d). In our example, $\hat{B} = \begin{bmatrix} 0 & 0 \\ 0 & .5 \end{bmatrix}$ ($R(\hat{B})$ is the x_2 axis) or $\hat{B} = \begin{bmatrix} .5 & 0 \\ 0 & 0 \end{bmatrix}$ ($R(\hat{B})$ is the x_1 axis) or infinitely

many other $\hat{B}$, as well as B^+, satisfy (A9) and (A10). The role of the constraints (A1.c) and (A1.d) is then to restrict R(A) to be identical with R(A*) thus obtaining the unique solution to (A1) which is the generalized inverse of A. This will be elaborated in the next section.

4. The Null Space of a Transformation

We shall first examine the concept of the "null space" of a transformation A, denoted by N(A), and then examine the relationship between R(A*) and N(A).

The null space of a transformation A is defined as a set of all vectors x in $\mathcal{X}_n$ which satisfy Ax = 0. For example, any vector in the null space of B satisfy $\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} x = 0$, or all the vectors in N(B) have a form $\binom{2}{-1}\alpha$ where α is a scalar. The vector $\binom{2}{-1}$ is a basis for the null space of B, or a null space basis of B. In the case of an m x n matrix A with rank r, the null space of A is n-r dimensional and therefore the null space basis consists of n-r linearly independent vectors. Of course, such a basis is not unique; (e.g. $\binom{1}{-.5}$, $\binom{4}{-2}$, etc., can also be a basis for N(B).) However, once we obtain a basis, any vector in N(A) can be expressed by a linear combination of the basis vectors.

A null space basis may be obtained by diagonalizing the matrix by an elementary row operation. For an m x n matrix A with rank r, we first obtain, by suitable row and column interchanges, if necessary, a matrix D which has a form $\begin{bmatrix} I_r & \vdots & H \\ \cdots & 0 & \cdots \end{bmatrix}$ where I_r is the r-dimensional identity matrix, 0 is the (m-r) x n zero matrix, and H is an r x (n-r) matrix of residuals, and has a property that A = EDP where E is a non-singular matrix of elementary row operation and P is a permutation matrix for column interchange. (Note that

the first r columns of D*, i.e. $\begin{bmatrix} I_r \\ H^* \end{bmatrix}$ constitute a basis for R(A*) since $A^* = (EDP)^* = P^*D^*E^*$ in which P* is simply a (non-singular) permutation matrix and E* is also a non-singular matrix. I.e. any vector in R(A*) can be represented by a linear combination of columns of $\begin{bmatrix} I_r \\ \cdots \\ H^* \end{bmatrix}$. Then, we prepare a matrix $\begin{bmatrix} -H \\ \cdots \\ I_{n-r} \end{bmatrix}$ and interchange <u>rows</u> for any corresponding <u>column</u> interchanges made during the diagonalization process. The resulting matrix, denoted by A^o, which is equal to $P^*\begin{bmatrix} -H \\ \cdots \\ I_{n-r} \end{bmatrix}$ where P* is the transpose of P, is a matrix for the null space basis of A, since $AA^o = EDPP^*\begin{bmatrix} -H \\ \cdots \\ I_{n-r} \end{bmatrix}$ $= E\left[\begin{array}{c:c} I_r & H \\ \hdashline & 0 \end{array}\right] \begin{bmatrix} -H \\ \cdots \\ I_{n-r} \end{bmatrix} = E0 = 0$. (See Orden, 1962.) For example, a 3 x 4 matrix given below may be diagonalized as follows to obtain a matrix D:

$$\begin{bmatrix} 2 & 2 & 4 & 2 \\ 4 & 4 & 8 & 4 \\ 3 & 3 & 7 & 5 \end{bmatrix} \dashrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \dashrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \dashrightarrow \left[\begin{array}{cc:cc} 1 & 0 & 1.5 & 1 \\ 0 & 1 & .5 & 0 \\ \hdashline 0 & 0 & 0 & 0 \end{array}\right]$$

Divide Row 1 by 2. Subtract 2 times Row 1 from Row 2. Subtract 3/2 times Row 1 from Row 3.

Interchange Row 2 and 3

Interchange Column 2 and 4

Divide Row 2 by 2. Subtract Row 2 from Row 1.

We then obtain a matrix $\begin{bmatrix} -H \\ I_{n-r} \end{bmatrix}$ which is $\begin{bmatrix} -1.5 & -1 \\ -.5 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. However, since we interchanged 2nd and 4th columns during the process of diagonalization, we interchange 2nd and 4th rows of $\begin{bmatrix} -H \\ I_{n-r} \end{bmatrix}$, obtaining A^o which is $\begin{bmatrix} -1.5 & -1 \\ 0 & 1 \\ 1 & 0 \\ -.5 & 0 \end{bmatrix}$. Note that any row interchange during the diagonalization process does not affect A^o. It may be easily checked that $AA^o = 0$.

Any vector in N(A) can then be repressed by $A^o z$ ($z \in \mathfrak{Z}_{n-r}$).

i.e., $N(A) = \{A^{O}z: \ z \in \mathcal{Z}_{n-r}\}$ On the other hand, $R(A^*)$ can be expressed by A^*w ($w \in \mathcal{Y}_m$) by the definition of $R(A^*)$. Then the inner product of any vector in $R(A^*)$ and any vector in $N(A)$ is always equal to zero since the inner product of A^*w and $A^{O}z$ is

(A11) $$(A^*w, A^{O}z) = w^*AA^{O}z = 0.$$

Therefore, the two subspaces $R(A^*)$ and $N(A)$ of $\mathcal{X}_n$ are said to be "orthogonal" (which means perpendicular in real vector spaces) with each other. (Refer to Halmos, 1958, p.122ff.)

We also define a "direct sum" which we associate with an operator $\oplus$ as follows:

If V_1 and V_2 are subspaces of a space V with the properties

(A12) $$V = V_1 + V_2 = \{x+y \mid x \in V_1, y \in V_2\}$$

(A13) $$V_1 \cap V_2 = 0$$

then $V = V_1 \oplus V_2$. Furthermore, if V_1 and V_2 are orthogonal with each other the two subspaces are called orthogonal complements of each other in the space V. (Note that for any given V_1 in V its orthogonal complement V_2 in unique.) In our case $N(A)$ is a subspace of $\mathcal{X}_n$ with dimension n-r and $R(A^*)$ is its orthogonal complement of dimension r, i.e.

(A14) $$\mathcal{X}_n = R(A^*) \oplus N(A),$$

since any basis of $R(A^*)$, e.g. $\begin{bmatrix} I_r \\ H^* \end{bmatrix}$, and any basis of $N(A)$, e.g. $\begin{bmatrix} -H \\ I_{n-r} \end{bmatrix}$ are linearly independent and collectively constitute n linearly independent vectors which span $\mathcal{X}_n$. In exactly the same way, we can show that

(A15) $\mathcal{Y}_m = R(A) + N(A^*)$.

In our example, $R(B^*)$ is a set of all vectors which has a form $\binom{1}{2}\alpha$ where α is a scalar and $N(B)$ is a set of all vectors which has a form $\binom{2}{-1}\beta$ where β is a scalar. Clearly, the inner product $(\alpha, 2\alpha)\binom{2\beta}{-\beta} = 0$ for any choice of α and β, and hence $R(B^*)$ and $N(B)$ are orthogonal. Moreover, since $\binom{1}{2}$ and $\binom{2}{-1}$ are two linearly independent vectors, any vector in $\mathcal{X}_2$ can be uniquely expressed by a sum of a vector in $R(B^*)$ and a vector in $N(B)$, i.e. $\mathcal{X}_2 = R(B^*) \oplus N(B)$. In the following diagram for $\mathcal{X}_2$, $R(B^*)$ and $N(B)$ are represented by the two lines which intersect prependicularly with each other at the origin.

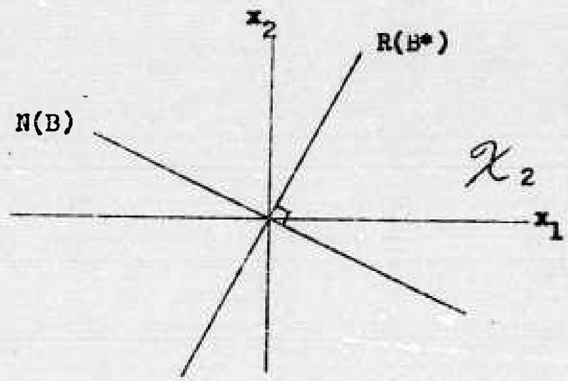


Figure A-1
R(B*) and N(B)

Similarly, $R(B)$ whose vectors are all in the form of $\binom{2}{1}\alpha$ and $N(B^*)$ whose vectors are all in the form of $\binom{-1}{2}\beta$ are represented in the following diagram for $\mathcal{Y}_2$ by the two lines which intersect perpendicularly with each other at the origin.

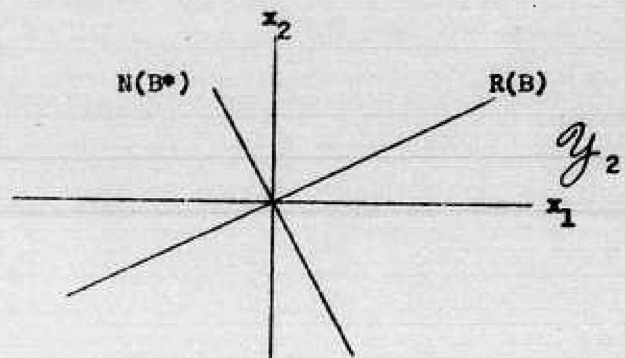


Figure A-2
R(B) and N(B*)

Now let us go back to the n x m matrix $\hat{A}$ which satisfies (A9) and (A10) and whose range, $R(\hat{A})$, has r basis vectors, none of which are transformed into the null vector in $\mathcal{Y}_m$ under the transformation A. The latter condition is equivalent to saying that any basis of $R(\hat{A})$ is linearly independent of any basis of N(A) and hence $\mathcal{X}_n = R(\hat{A}) \oplus N(A)$. Therefore, any vector x in $\mathcal{X}_n$ can be uniquely represented by the sum of the vector in $R(\hat{A})$ which is obtained by projecting x on $R(\hat{A})$ along with (or parallel to) N(A) and the vector in N(A) which is obtained by projecting x on N(A) along with (or parallel to) $R(\hat{A})$. For example, if we set $\hat{B} = \begin{bmatrix} 0 & 0 \\ 0 & .5 \end{bmatrix}$, $R(\hat{B})$ is all the points on x_2 axis. Then, as shown in the following diagram, any point x (e.g., $\binom{6}{2}$) in $\mathcal{X}_2$ can be represented by the sum of the vector x_1 (e.g., $\binom{0}{5}$) in $R(\hat{B})$ which is obtained by projecting x on $R(\hat{B})$ along with N(B) and the vector

x_2 (e.g., $\binom{6}{-3}$) in N(B) which is obtained by projecting x on N(B) along with $R(\hat{B})$.

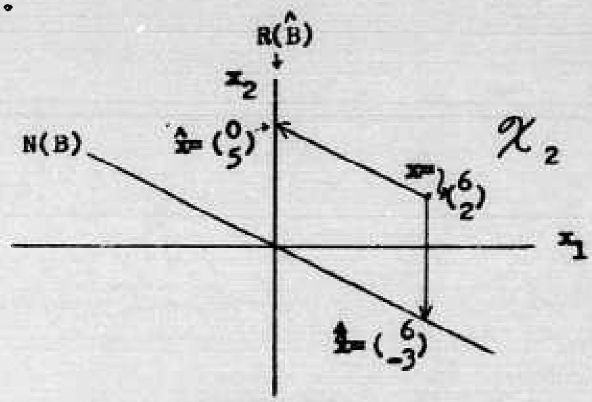


Figure A-3
Projection on $R(\hat{B})$ and N(B)

Next we shall show that for any x the vector $\hat{x} = \hat{A}Ax$ is the vector in $R(\hat{A})$ projected from x parallel to N(A). First, note that the set of all vectors which correspond to a given vector $\hat{y}$ in $\mathcal{Y}_m$ under the transformation A constitute a subspace of $\mathcal{X}_n$ which is parallel to N(A) since if not any point in the intersection of the two subspaces corresponds to a given vector $\hat{y}$ and at the same time to the null vector which is contradictory (unless $\hat{y} \equiv 0$ in which case the two subspaces are identical.) Since $A\hat{x} = A\hat{A}Ax = Ax$ by (A10), we conclude that both x and $\hat{x} = \hat{A}Ax$ correspond to the same vector in $\mathcal{Y}_m$, and hence $\hat{x}$ is the vector obtained from x by projecting on $R(\hat{A})$ parallel to N(A). The vector $\hat{\hat{x}}$ in N(A) which is obtained by projecting x on N(A) parallel to $R(\hat{A})$ is given simply by $x - \hat{x} = x - \hat{A}Ax$.

We then seek for a condition on A that makes the two subspaces $R(\hat{A})$ and $N(A)$ orthogonal. In order for this to be true we have to have the inner product of $\hat{x}$ and $\hat{\hat{x}}$ equal to zero for any x in $\mathcal{X}_n$. I.e.

(A16) $$(\hat{A}Ax, x-\hat{A}Ax) = x^*(\hat{A}A)^*x - x^*(\hat{A}A)^*\hat{A}Ax$$
$$= x^*((\hat{A}A)^* - (\hat{A}A)^*(\hat{A}A))x = 0 \quad \text{for all } x \in \mathcal{X}_n,$$

or

(A17) $$(\hat{A}A)^* = (\hat{A}A)^*(\hat{A}A)$$

Let $(\hat{A}A)^* = \hat{A}A + X$. Then, (A17) is rewritten as follows:

(A18) $$\hat{A}A + X = (\hat{A}A + X)(\hat{A}A) = \hat{A}A\hat{A}A + XAA = \hat{A}A + X\hat{A}A,$$

or

(A19) $$\hat{A}AX = X$$

This means that we have to have either $\hat{A}A = I$ or $X = 0$ which implies $\hat{A}A = (\hat{A}A)^*$. Since if $\hat{A}A = I$, $(\hat{A}A)^* = \hat{A}A$ which implies $X = 0$, the former case is simply a special case of the latter. Therefore, we conclude that the necessary and sufficient condition for $R(\hat{A})$ and $N(A)$ to be orthogonal is

(A20) $$\hat{A}A = (\hat{A}A)^*$$

In exactly the same way we obtain the necessary and sufficient condition for $R(A)$ and $N(\hat{A})$ to be orthogonal as

(A21) $$A\hat{A} = (A\hat{A})^*$$

Note that (A20) and (A21) are exactly the same as (A1.c) and (A1.d) used in defining the generalized inverse of A. Therefore, we have shown that $R(A^\dagger)$ is the orthogonal complementary subspace to $N(A)$ in $\mathcal{X}_n$ and that by the uniqueness of the orthogonal complementary subspace,

$R(A^\dagger) = R(A^*)$.

In summary, $A^\dagger$ is a transformation which transforms vectors in $\mathcal{Y}_m$ into $R(A^*)$ which is a subspace of $\mathcal{X}_n$ in such a way that for any vector $\hat{x}$ in $R(A^*)$ $A^\dagger A\hat{x} = \hat{x}$, and for any vector $\hat{y}$ in $R(A)$ $AA^\dagger\hat{y} = \hat{y}$. In particular, if A is an n-dimensional non-singular matrix, $R(A) \equiv \mathcal{Y}_n$ and $R(A^*) \equiv \mathcal{X}_n$ and hence for any vector x in $\mathcal{X}_n$ $A^\dagger Ax = x$ and for any vector y in $\mathcal{Y}_n$ $AA^\dagger y = y$. However, since this is true if and only if $A^\dagger A = AA^\dagger = I$, we have $A^\dagger = A^{-1}$ when A is non-singular. This means that a generalized inverse has the same inverting property as an ordinary inverse has but the objects to which the inverting property works are limited to the range of the original matrix and the range of the transposed matrix in the case of a generalized inverse.

5. Solutions to a Linear System

By the use of a generalized inverse solutions to a linear system $Ax = y$ can now be obtained easily. Let us first consider a linear system

(A22) $$Ax = \hat{y}$$

where $\hat{y}$ is a given vector in $R(A)$. Then a solution $\hat{x}$ to (A22) is simply

(A23) $$\hat{x} = A^{\dagger}\hat{y},$$

since $A\hat{x} = AA^{\dagger}\hat{y} = \hat{y}$ by the inverting property of a generalized inverse. However, this is only one of many possible solutions to $Ax = \hat{y}$. Note that if $\hat{x}$ is a solution to $Ax = \hat{y}$, any vector x which is a sum of $\hat{x}$ and any vector x° in $N(A)$ is also a solution to $Ax = \hat{y}$, for $Ax = A\hat{x} + Ax^{\circ} = A\hat{x} = \hat{y}$. Furthermore x° has to be a vector in $N(A)$, otherwise $Ax = A\hat{x} + Ax^{\circ} \neq A\hat{x} = \hat{y}$. Therefore, the set of <u>all</u> solutions to $Ax = \hat{y}$ can be obtained by adding each vector in $N(A)$ to $\hat{x}$. Thus the uniqueness of a solution to $Ax = \hat{y}$ depends entirely upon whether or not $N(A)$ consists of only the null vector which is true if and only if A is non-singular. Since any vector in $N(A)$ may be represented by $A^{\circ}z$ $z \in \mathcal{Z}_{n-r}$, the set of all solutions to $Ax = \hat{y}$, denoted by $S(A, \hat{y})$ is given by:

(A24) $$S(A, \hat{y}) = \left\{ x = A^{\dagger}y + A^{\circ}z : \ z \in \mathcal{Z}_{n-r} \right\}.$$

The use of the matrix A° of the null space basis is not the only way to represent $N(A)$. We shall examine another way of representing $N(A)$ by means of $A^{\dagger}$. Take any vector x in $\mathcal{X}_n$. Then Ax generates a vector $\hat{y}$ in $R(A)$, from which we obtain $A^{\dagger}Ax$ by transforming back

to $R(A^*)$ by means of $A^\dagger$. The vector x thus obtained is a projection of x on $R(A^*)$ parallel to $N(A)$ as discussed in the previous section. Since $R(A^*)$ and $N(A)$ are orthogonal (or perpendicular), x may be considered as an orthogonal projection of x on $R(A^*)$. Then $x - A^\dagger Ax$ or $(I - A^\dagger A)x$ gives us a vector in $N(A)$ which is obtained by an (orthogonal) projection of x on $N(A)$ along with $R(A^*)$. By allowing x to vary over $\mathcal{X}_n$ we can obtain every vector in $N(A)$ by the expression $(I - A^\dagger A)x, x \in \mathcal{X}_n$. Thus (A24) may be rewritten as

(A25) $$S(A, \hat{y}) = \{ x = A^\dagger y + (I - A^\dagger A)z: \; z \in \mathcal{X}_n \} .$$

For example, if we set $\hat{y} = \binom{10}{5}$ in $Bx = \hat{y}$, we obtain a solution $\hat{x}$ in $R(B)$ by (A23), i.e. $\begin{bmatrix} .08 & .04 \\ .16 & .08 \end{bmatrix}\binom{10}{5} = \binom{1}{2}$. Since vectors in $N(B)$ are all in the form of $\binom{2}{-1}\alpha$ as we have observed, the set of all solutions to $Bx = \binom{10}{5}$ is given by

(A26) $$x = \binom{1}{2} + \binom{2}{-1}\alpha \quad (\alpha \text{ being a scalar})$$

In Figure A-4 below, the set of all solutions to $Bx = \binom{10}{5}$, or $S(B, \binom{10}{5})$, is represented by a line which is parallel to the line for $N(B)$.

The solution obtained by $A^\dagger\hat{y}$ has a special property which does not exist in any other solutions to $Ax = \hat{y}$. Notice that since $S(A, \hat{y})$ is a parallel shift of $N(A)$ which is orthogonal to $R(A^*)$, $S(A, \hat{y})$ is orthogonal to $R(A^*)$. This implies that among all the vectors in $S(A, \hat{y})$ the vector at the intersection of $S(A, \hat{y})$ and $R(A^*)$, which is $A^\dagger\hat{y}$, has a minimum Euclidean distance to the origin, i.e. the square root of (x, x), or equivalently, $(x_1^2 + x_2^2 + \ldots x_n^2)^{1/2}$, where $x_1, x_2, \ldots, x_n$ are the elements of x, is minimum among all vectors

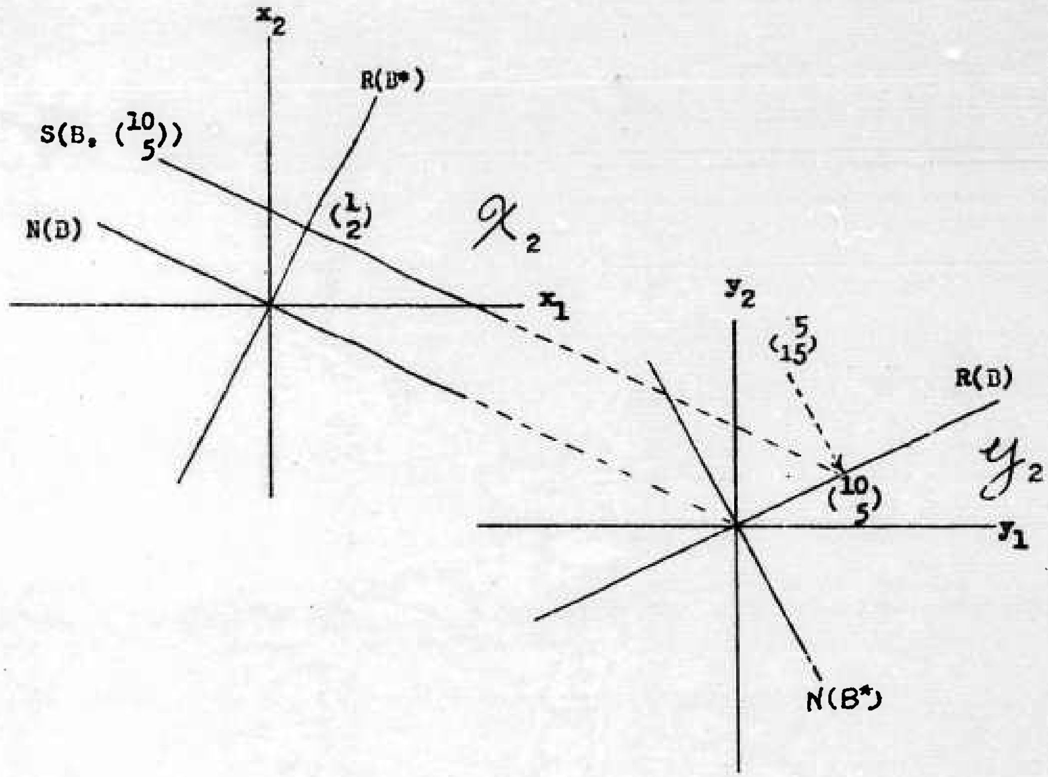


Figure A-4

The Solution Space of

$Bx = \binom{10}{5}$

x in $S(A, y)$. This quantity, the square root of (x, x), is called the norm of x and is denoted by $||x||$. In our example, the norm of $A^{+}b$ $(= \binom{1}{2})$ is $\sqrt{1^2 + 2^2} = \sqrt{5}$. This norm is minimum among all vectors in $S(B, \binom{10}{5})$ which are in the form of $\binom{1 + 2\alpha}{2 - \alpha}$ (as given in (A26)) where α is a scalar, since clearly $[(1 + 2\alpha)^2 + (2 - \alpha)^2]^{1/2}$ $= [5\alpha^2 + 5]^{1/2}$ is minimum when $\alpha = 0$.

Therefore, we can restore the uniqueness of a solution to $Ax = y$ by requiring the norm of the solution to be minimum. Since for any subspace of a vector space there is one and only one vector which has a minimum Euchlidean distance to the origin we can always find a unique solution to such a minimizing problem provided that $S(A, \hat{y})$ is non-empty, or equivalently, $\hat{y}$ is in $R(A)$.

In summary, $A^{+}b$ is the unique solution to

(A27) $$\text{Minimize } (x, x)^{1/2} \quad \text{Subject to } Ax = \hat{y},$$

or, equivalently,

(A28) $$\text{Minimize } (x, x) \quad \text{Subject to } Ax = \hat{y}.$$

Having used this minimizing property of A^{+} to obtain the "uniqueness" that we want to associate with our concept of "solution", we next turn to its use in connection with the question of "existence."

Consider a linear system $Ax = y$ in which y does not lie in $R(A)$. In such a case, no solution exists for $Ax = y$ by the definition of $R(A)$. What, then, is the significance of $A^{+}y$ when y is not in $R(A)$? Note that $AA^{+}y$ is the orthogonal

projection of y on R(A) just as $A^\dagger Ax$ is the orthogonal projection of x on $R(A^*)$ which we have already examined. Since $AA^\dagger y$ is in R(A), solutions exist to $Ax = AA^\dagger y$ which is given by

(A29) $$S(A, AA^\dagger y) = \{x = A^\dagger AA^\dagger y + (I - A^\dagger A)z: \; z \in \mathcal{X}_n\},$$

using the formula given by (A25). However, since $A^\dagger AA^\dagger = A^\dagger$ by (A1.b), (A29) is equivalent to

(A30) $$S(A, AA^\dagger y) = \{x = A^\dagger y + (I - A^\dagger A)z: \; z \in \mathcal{X}_n\}.$$

Let us then examine the significance of the vector $\hat{y} = AA^\dagger y$ in its relation with y. Since $A^\dagger\hat{y} = A^\dagger AA^\dagger y = A^\dagger y$, both $\hat{y}$ and y lie in the same solution space to $A^\dagger y = \hat{x}$ ($\hat{x} \in R(A^*)$). However, a set of all solutions to $A^\dagger y = \hat{x}$ ($\hat{x} \in R(A^*)$) constitute a solution space $S(A^\dagger, x)$ which is parallel to $N(A^\dagger)$ just as in the case of solutions to $Ax = \hat{y}$ ($\hat{y} \in R(A)$).

Since $N(A^\dagger)$ is orthogonal to R(A), so is $S(A^\dagger, \hat{x})$. This means that $\hat{y}$ is a vector obtained by projecting y orthogonally to R(A), and hence $\hat{y}$ is a vector in R(A) which has a minimum Euclidean distance to y among all vectors in R(A). (E.g., $y=(1,\xi)$ in Figure A-4). Therefore, a set of all vectors given by

(A31) $$x = A^\dagger y + (I - A^\dagger A)z \quad (z \in \mathcal{X}_n)$$

is the solution set to $Ax = AA^\dagger y = \hat{y}$ where $\hat{y}$ is a vector in R(A) which has a minimum Euclidean distance to y. In particular if $y = \hat{y}$ or, in other words, if y is in R(A), the minimum Euclidean distance is zero. In other words, (A31) gives us the set of all solutions to the problem;

(A32) $$\text{Minimize} \quad ||Ax - y||,$$

where $||Ax - y||$ is the norm of the vector $(Ax - y)$, i.e. $(Ax - y, Ax - y)^{1/2}$. In particular, if it is possible to have an x which brings $||Ax - y||$ down to zero, $Ax = y$ is solvable and the set of all solutions is given by (A31). Thus, the original problem of $Ax = y$ which may or may not have a solution is changed into a problem of minimizing the Euclidean distance between Ax and y, or a problem of finding x which makes Ax as "close to y" as possible, for which the existence of a solution is guaranteed since for any vector y in $\mathcal{Y}_m$ there exists (uniquely) a vector y in R(A) which minimizes the Euclidean distance to y so far as R(A) is non-empty which is true even for $A = 0$.

In summary, $A^{\dagger}y$ is the <u>unique</u> solution, which is guaranteed to <u>exist</u>, to the following double minimization problem:

(A33) Minimize $||x||$

Subject to $||Ax - y||$ to be minimum,

where x is a variable in $\mathcal{X}_n$ and y is a given vector in $\mathcal{Y}_m$.

6. An Example of Computation of $A^\dagger$

In Appendix B, various methods of calculating generalized inverses are described, together with a flow chart and GATE program for one of these methods. In this section, however, we shall show the process of calculating a generalized inverse by an "elimination method" by using an example. For more detailed discussions on this method refer to Appendix B.

Let us calculate the generalized inverse of a 3 x 4 matrix $\begin{bmatrix} 2 & 2 & 4 & 2 \\ 4 & 4 & 8 & 4 \\ 3 & 3 & 7 & 5 \end{bmatrix}$. If the number of rows in a matrix is less than the number of columns, it is easier to calculate the generalized inverse of the transposed matrix and then transpose the resulting matrix, obtaining the generalized inverse of the original matrix, since $(A^*)^\dagger = (A^\dagger)^*$ which is clear from our earlier discussions.

Let us therefore in this case consider A to be

$$A = \begin{bmatrix} 2 & 4 & 3 \\ 2 & 4 & 3 \\ 4 & 8 & 7 \\ 2 & 4 & 5 \end{bmatrix}$$

The first step is to calculate A^*A.

$$A^*A = \begin{bmatrix} 28 & 56 & 50 \\ 56 & 112 & 100 \\ 50 & 100 & 92 \end{bmatrix}.$$

Then, form a matrix $[A^*A \mid A^*]$ and diagonalize by a set of elementary row operations, interchanging rows and columns if necessary. (Columns of A^* should never be interchanged with other columns in A^*A and A^*.)

The diagonalization process is as follows. We shall denote an element in i^{th} row and j^{th} column of the tableau by b_{ij}.

$$\left[\begin{array}{ccc|cccc} 28 & 56 & 50 & 2 & 2 & 4 & 2 \\ 56 & 112 & 100 & 4 & 4 & 8 & 4 \\ 50 & 100 & 92 & 3 & 3 & 7 & 5 \end{array}\right] \quad [A^*A \mid A^*]$$

$$\left[\begin{array}{ccc|cccc} 1 & 2 & \frac{50}{28} & \frac{2}{28} & \frac{2}{28} & \frac{4}{28} & \frac{2}{28} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{76}{28} & -\frac{16}{28} & -\frac{16}{28} & -\frac{4}{28} & \frac{40}{28} \end{array}\right]$$

Divide Row 1 by b_{11} and subtract scalar multiples of Row 1 from Row 2 and Row 3, the scalars being given by b_{21} and b_{31}.

$$\left[\begin{array}{ccc|cccc} 1 & \frac{50}{28} & 2 & \frac{2}{28} & \frac{2}{28} & \frac{4}{28} & \frac{2}{28} \\ 0 & \frac{76}{28} & 0 & -\frac{16}{28} & -\frac{16}{28} & -\frac{4}{28} & \frac{40}{28} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Interchange Row 2 and Row 3, then Column 2 and Column 3 to bring a non-zero element in b_{22}.

$$\left[\begin{array}{ccc|cccc} 1 & 0 & 2 & \frac{34}{76} & \frac{34}{76} & \frac{18}{76} & -\frac{66}{76} \\ 0 & 1 & 0 & -\frac{16}{76} & -\frac{16}{76} & -\frac{4}{76} & \frac{40}{76} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Divide Row 2 by b_{22} and subtract scalar multiples of Row 2 from Row 1 and Row 3, the scalars being given by b_{12} and b_{32}.

We have diagonalized the first two rows but we cannot procced the diagonalization further since the remaining row contains only zeros. This ends the diagonalization process. The result of diagonalization is in the form of $\left[\begin{array}{c:c} I_r \; \Delta & EA^* \\ \hdashline 0 & \end{array}\right]$, where I_r is the r-dimensional identity matrix, Δ is an $r \times (n-r)$ matrix of residuals, 0 is the $(n-r) \times n$ zero matrix, and E is an $n \times n$ matrix of elementary row operations. The dimension of I_r, i.e., r, is the rank of A^*A which is equal to the rank of A. Let us denote the first r rows of EA^* by $(EA^*)_r$.

The next step is, then, to calculate $(I_r + \Delta\Delta^*)^{-1}$, which is guaranteed to exist for any Δ. In our example,

$$(I_r + \Delta\Delta^*)^{-1} = \left[\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \begin{pmatrix}2\\ 0\end{pmatrix}(2,\ 0)\right]^{-1} = \begin{pmatrix}5 & 0\\ 0 & 1\end{pmatrix}^{-1} = \begin{pmatrix}.2 & 0\\ 0 & 1\end{pmatrix}.$$

We then calculate $\begin{bmatrix}I_r\\ \Delta^*\end{bmatrix}(I_r + \Delta\Delta^*)^{-1}(EA^*)_r$. In our example,

$$\begin{bmatrix}1 & 0\\ 0 & 1\\ 2 & 0\end{bmatrix}\left[\frac{1}{5}\begin{pmatrix}1 & 0\\ 0 & 5\end{pmatrix}\right]\left[\frac{1}{76}\begin{pmatrix}34 & 34 & 18 & -66\\ -16 & -16 & -4 & 40\end{pmatrix}\right]$$

$$= \frac{1}{380}\begin{bmatrix}34 & 34 & 18 & -66\\ -80 & -80 & -20 & 200\\ 68 & 68 & 36 & -132\end{bmatrix} = \frac{1}{190}\begin{bmatrix}17 & 17 & 9 & -33\\ -40 & -40 & -10 & 100\\ 34 & 34 & 18 & -66\end{bmatrix}$$

Fianlly we permute rows of the resulting matrix if we interchanged corresponding columns during the diagonalization process. (Any row interchange during diagonalization does not affect the result and hence can be ignored.) In our example, since we interchanged the second and the third columns. Hence we now interchange the second and the third rows of the above matrix, obtaining the generalized inverse of A,

$$A^+ = \frac{1}{190}\begin{bmatrix}17 & 17 & 9 & -33\\ 34 & 34 & 18 & -66\\ -40 & -40 & -10 & 100\end{bmatrix}$$

If A is the transpose of the original matrix, we transpose A^+ to obtain the generalized inverse of the original matrix. Thus,

$$\begin{bmatrix}2 & 2 & 4 & 2\\ 4 & 4 & 8 & 4\\ 3 & 3 & 7 & 5\end{bmatrix}^+ = \frac{1}{190}\begin{bmatrix}17 & 34 & -40\\ 17 & 34 & -40\\ 9 & 18 & -10\\ -33 & -66 & 100\end{bmatrix}$$

It may be verified that (A1.a) - (A1.d) are all satisfied by the matrix thus calculated.

If the i^{th} row of A is all zeros, this row may be eliminated before we start the calculation. We then insert a zero column

between the (i-1)st column and the i^{th} column of $A^{\dagger}$ obtained at the end of cðluulation. Similarly, if j^{th} column of A is all zeros, this column may be eliminated before we start the calculation, inserting a zero row between (j-1)-st row and the j^{th} row of $A^{\dagger}$ obtained at the end of the calculation.

7. A List of Some Properties of Generalized Inverses

The generalized inverse has the following properties. (See Penrose, 1955; Cline, 1962.)

1. $A = 0$ (m by n) implies $A^\dagger = 0$ (n by m);
2. $A^{\dagger\dagger} = A$;
3. $A^{*\dagger} = A^{\dagger *}$;
4. if A is non-singular $A^\dagger = A^{-1}$,
5. $(A^*A)^\dagger \; A^\dagger A^{\dagger *}$;
6. if U and V are unitary $(UAV)^* = V^*A^\dagger U^*$;
7. if $A = \sum A_i$, where $A_iA_j^* = 0$ and $A_i^*A_j = 0$, whenever $i \neq j$, then $A^\dagger = \sum A_i^\dagger$;
8. if A is normal $A^\dagger A = AA^\dagger$ and $(A^n)^+ = (A^\dagger)^n$;
9. A, A^*A, A^* and $A^\dagger A$ all have rank equal to trace $A^\dagger A$;
10. $A^\dagger A$, $AA^\dagger$, $I-A^\dagger A$, and $I-AA^\dagger$ are all hermitian and idempotent ;
11. $(\lambda A)^\dagger = \lambda^\dagger A^\dagger$, whera λ is a complex number and $\lambda^\dagger$ means λ^{-1} if $\lambda \neq 0$ and 0 if $\lambda = 0$;
12. if A is hermitian and idempotent $A^\dagger = A$;
13. if A has full column rank $A^\dagger = (A^*A)^{-1}A^*$;
14. if A has full row rank $A^\dagger = A^*(AA^*)^{-1}$;
15. if B (m by r), C (r by r), and D (r by n) each has rank r, where $1 \leq r \leq$ minimum (m, n), $(BCD)^\dagger = D^\dagger C^\dagger B^\dagger$.

APPENDIX B

A REPORT ON THE MACHINE COMPUTATION
OF THE
GENERALIZED INVERSE OF AN ARBITRARY MATRIX[1]

1. The Generalized Inverse of an Arbitrary Matrix

E. H. Moore [27, 28] and independently Bjerhammar [6] and Penrose [30] showed that the concept of inverses can be generalized as follows:

For any m x n complex matrix A, the system of matric equations

$$AXA = A$$
$$XAX = X$$
$$(AX)^* = AX \qquad \text{(* denotes: conjugate transpose)}$$
$$(XA)^* = XA$$

has a unique solution, an n x m complex matrix called the generalized inverse of A, denoted by A^+.

If A is nonsingular then $A^+ = A^{-1}$, otherwise A^+ still possesses many desirable properties which make it a central concept in linear algebra, numerical analysis and in various applications.

The main references on generalized inverses are: Moore [27, 28], Bjerhammar [6, 7, 8], Penrose [30], Greville [16], Hestenes [19, 20], Lanczos [25], Rado [36], den Broeder-Charnes [9] and Ben-Israel-Charnes [2]. A representative list of references on applications of generalized inverses is: [3], [6], [7], [11], [16], [17], [22], [23], [24], [31], [32], [35], [42] and [43].

---

[1] Reproduced from A. Ben-Israel and Y. Ijiri, "A Report on the Machine Computation of the Generalized Inverse of an Arbitrary Matrix," ONR Research Memorandum No.110. Pittsburgh: Carnegie Institute of Technology, March, 1963. All the references in this appendix are give on page B-15-B-17.

2. Computational Methods

If A is nonsingular ($A^+ = A^{-1}$) or of full rank ($A^+ = (A^*A)^{-1}A^*$ or $A^+ = A^*(AA^*)^{-1}$) then A^+ may be computed by ordinary inversion methods, e.g. [13], [21] and [27]. If A is not of full rank, then a full rank matrix $\tilde{A}$ may be formed by adjoining to A sufficiently many rows (columns) (orthogonal to those of A) and A^+ is obtained from $\tilde{A}^+$ by striking out the corresponding columns (rows), e.g. [18]. Methods for computing A^+, when A is of arbitrary rank, are not necessarily trivial extensions of the ordinary inversion algorithms — although they reduce to the later algorithms when A is nonsingular. A review of some of these methods and of some other results of computation interest will now be given:

2.1 Direct Methods:

2.1.1 Elimination Methods

2.1.11 If S is a symmetric matrix

C a nonsingular matrix (product of elementary matrices) such that

$CSC^* = D = (d_{ii})$ diagonal

then

$S^+ = C^*D^+C$ where $D^+ = (d_{ii}^{\;+})$ (Rao [35])

2.1.12 Similarly if A is a square matrix, V, W unitary matrices[1/]

such that

$A = VDW, \qquad D = (d_{ii})$

then

$A^+ = W^*D^+V^*$ (Penrose [30]

1/ Existence guaranteed by a theorem of Autonne, extended to the rectangular case in [18].

2.1.13 For an $m \times n$ matrix A, let E be a nonsingular matrix (a product of elementary matrices) and P be a permutation matrix such that

$$EA^*AP = \left(\begin{array}{c:c} I_r & \Delta \\ \hdashline \multicolumn{2}{c}{0} \end{array}\right) \qquad \text{where } r = \text{rank } A$$

and Δ is determined by the particular E, P. Then

$$A^+ = P\left(\begin{array}{c} I_r \\ \hdashline \Delta^* \end{array}\right) \left((I_r + \Delta\Delta^*)^{-1} \;\vdots\; 0\right)EA^* \qquad ([10])$$

2.1.14 With E, P, Δ as above let $D = \left(\begin{array}{c} \Delta \\ \hdashline -I_{n-r} \end{array}\right)$. Then

$$A^+ = P(I_n - D(D^*D)^{-1}D^*)EA^* \qquad ([4]).$$

2.1.2 Orthogonalization Methods

2.1.21 Hestenes' biorthogonalization method [18], essentially equivalent to one form of the Gauss' elimination method[1/], can be used to compute A^+ in the case where A is of full rank, ibid. p. 60.

2.1.22 Other orthogonalization methods, e.g. [2] pp. 72-78, may be modified for use in the computation of A^+.

2.1.3 Escalator and Partitioning Methods

2.1.31 Let A_k, the matrix composed of the first k columns of A: $a_1, \ldots, a_k$, be partitioned as

$$A_k = (A_{k-1} \;\vdots\; a_k)$$

then

$$A_k^+ = \left(\begin{array}{c} A_{k-1}^+ - d_k b_k^* \\ \hdashline b_k^* \end{array}\right) \qquad k = 2, 3, \ldots, n$$

1/ Ibid. pp. 69-71.

where $\quad d_k = A_{k-1}^{+} a_k$

$$b_k = (1 + d_k^* d_k)^{-1} d_k^* A_{k-1}^{+}$$

This method due to Greville [17] is a recursive procedure for obtaining A_k^+ from A_{k-1}^+. See also [40], and a generalization in [12].

<u>2.1.32</u> If A is partitioned into proper submatrices, then expressions for the corresponding submatrices in A^+ are given in [31] and in [6]. Expressions for A^+ in terms of the subdeterminants of A, A^* were originally given by Moore [27], see also [37, 38].

<u>2.1.4 Frame's Method</u>

Penrose [31] adapted Frame's method (e.g. [13]) for computing A^+ as follows:

Let $\quad C_{(1)} = I$

$$C_{(j+1)} = \left(\frac{1}{j} \operatorname{trace}\left\{ C_{(j)} A^* A \right\}\right) I - C_{(j)} A^* A \qquad j = 2, \ldots, r$$

(The process terminates for $j = r$ because $C_{(r)} A^* A = 0$.)

and let

$$D = \frac{r C_{(r)}}{\operatorname{trace}\{C_{(r)} A^* A\}}$$

then

$$A^+ = DA^*$$

<u>2.1.5 Inverting Modified Matrices</u>

den Broeder and Charnes [9] showed that if A is m x n and if $y^* A^+ x \neq 1$ then

$$(A - AA^+ x y^* A^+ A)^+ = A^+ + \frac{H^+ x y^* H^+}{1 - y^* H^+ x} \quad .$$

As in the nonsingular case, e.g. [21] p. 83 and [41], the above may be used in an inversion algorithm.

2.2 Approximate Methods

2.2.1 Iterative Methods

2.2.11 den Broeder and Charnes [9] showed that

$\lim_{n\to\infty} \sum_{k=1}^{n} A^*(I + AA^*)^{-k}$ always exists and

$$\sum_{k=1}^{\infty} A^*(I + AA^*)^{-k} = A^+$$

This result may be rewritten as:

If $\qquad X_0 = A^*(I + AA^*)^{-1}$

$$X_{n+1} = X_n(I + (I+AA^*)^{-1}) \qquad n = 0, 1, 2, \ldots$$

then

$$\lim_{n\to\infty} X_n = A^+$$

2.2.12 Ben-Israel and Charnes [2] showed that for any A

$\lim_{n\to\infty} \alpha \sum_{k=0}^{n} (I - \alpha A^*A)^k A^*$ always exists

and

$$\alpha \sum_{k=0}^{\infty} (I - \alpha A^*A)^k A^* = A^+$$

provided that

$$0 < \alpha < \frac{2}{\lambda_1} \qquad \lambda_1 \text{ is the greatest eigenvalue of } A^*A.$$

An iterative method of computing A^+ follows by rewriting the above as:

If $X_0 = \alpha A^*$ where $0 < \alpha < \frac{2}{\lambda_1}$

$$X_{n+1} = X_n + \alpha A^*(I - AX_n) \qquad n = 0, 1, 2, \ldots$$

then $\lim_{n\to\infty} X_n = A^+$

This is essentially a variant of the method of steepest descent, e.g. [21], pp. 49-52 and [14] p. 311. For related results see [33], [1] and [5].

2.2.2 Limit Method

If $\{\varepsilon_n\}$ is a sequence of positive numbers with $\varepsilon_n \to 0$, then

$$\lim_{n\to\infty} A^*(\varepsilon_n I + AA^*)^{-1} = A^+ \qquad \text{(den Broeder-Charnes [9])}$$

3. An Elimination Method for Computing A^+

In this section we present the basic steps of the elimination method 2.1.13 and relate these steps to the flow charts given in Appendix 1. A GATE-20 program GENINV is given in Appendix 2.

3.1 A Description of the Method

Step 1: Initialization

Starting with the m x n real matrix A (Flow Chart 1), strike out all the zero rows and the zero columns and record the corresponding indices (Flow Chart 2). A reduced matrix $\bar{A}$, of size $\bar{m} \times \bar{n}$ is thus obtained. Proceed to compute $\bar{A}^+$.

Step 2:

If $\bar{m} \geq \bar{n}$ set $\bar{\bar{m}} = \bar{m}$, $\bar{\bar{n}} = \bar{n}$ and $\bar{\bar{A}} = \bar{A}$

If $\bar{m} < \bar{n}$ set $\bar{\bar{m}} = \bar{n}$, $\bar{\bar{n}} = \bar{m}$ and $\bar{\bar{A}} = \bar{A}^*$

(Flow Chart 3)

Step 3:

Form $\bar{\bar{A}}^*\bar{\bar{A}}$ (Flow Chart 4)

Step 4: Diagonalization

Diagonalize the matrix $(\bar{\bar{A}}^*\bar{\bar{A}} \mid \bar{\bar{A}}^*)$ to obtain

$$(E\bar{\bar{A}}^*\bar{\bar{A}}P \mid E\bar{\bar{A}}^*) = \left(\begin{array}{c|c} I_r & \Delta \\ \hline 0 & 0 \end{array} \middle| \; E\bar{\bar{A}}^*\right)$$

(Flow Chart 5)[1/]

Record the permutation matrix P

the residue matrix Δ

the matrix $\underbrace{E\bar{\bar{A}}^*}$, consisting of the first r rows of $E\bar{\bar{A}}^*$,

and the rank r.

If $r = \bar{\bar{n}}$ (in which case $\Delta = 0$) go to step 9.[2/]

If $r < \bar{\bar{n}}$ go to step 5.

Step 5:

Form $(I_r + \Delta\Delta^*)$ (Flow Chart 5)

Step 6: Inversion

Invert $(I_r + \Delta\Delta^*)^{-1}$ (Flow Chart 6)

Step 7:

Form $\left(\frac{I_r}{\Delta^*}\right)(I_r + \Delta\Delta^*)^{-1}$ (Flow Chart 7)

1/ Because of convenience in machine operations we actually diagonalize $\left(\frac{\bar{\bar{A}}^*\bar{\bar{A}}}{\bar{\bar{A}}}\right)$ to obtain $\left(\begin{array}{c|c} I_r & 0 \\ \hline \Delta^* & 0 \\ \hline E\bar{\bar{A}} & \end{array}\right)$

2/ Which is simplified to: Form $PE\bar{\bar{A}}^* = \bar{\bar{A}}^+$.

Step 8:

Form $\binom{I_r}{\Delta^*}(I_r + \Delta\Delta^*)^{-1}\underbrace{\bar{\bar{EA}}^*}$ (Flow Chart 8)

Step 9: Permutation

Form $P\binom{I_r}{\Delta^*}(I_r + \Delta\Delta^*)^{-1}\underbrace{\bar{\bar{EA}}^*} = \bar{\bar{A}}^+$ (Flow Chart 9)

Step 10:

If $\bar{m} \geq \bar{n}$ then $\bar{A}^+ = \bar{\bar{A}}^+$

If $\bar{m} < \bar{n}$ transpose $\bar{\bar{A}}^+$ to obtain $\bar{A}^+ = (\bar{\bar{A}}^+)^*$ (Flow Chart 10)[1]

Step 11:

Insert zero rows (columns) in $\bar{A}^+$, in place of the zero columns (rows) stricken out in Step 1, to obtain A^+ (Flow Chart 11)

3.2 Operational Counts

3.2.1 In counting the arithmetic operations required by the above method (disregarding bookkeeping operations) we assume no zero rows or columns so that $A = \bar{A}$, $m = \bar{m}$ and $n = \bar{n}$. We assume also that $m \geq n$ and that $n > r$ so that Steps 5-8 are executed. The number of arithmetic operations in Step 6: the inversion of an r x r matrix, is left out as it depends on the method used for the inversion routine. The results are summarized below:

---

[1] Since the diagonalization data was written as $\left(\frac{\bar{\bar{A}}^*\bar{\bar{A}}}{\bar{\bar{A}}}\right)$, see Flow Chart 5, we actually obtain $(\bar{\bar{A}}^+)^*$ at Step 9. Therefore, according to Flow Chart 10, Step 10 is: If $\bar{m} \geq \bar{n}$ transpose $(\bar{\bar{A}}^+)^*$ to get $\bar{A}^+$.
If $\bar{m} < \bar{n}$ set $\bar{A}^+ = \bar{\bar{A}}^+$.

Step \ Operation	Division	Multiplication	Addition
3		mn^2	mn^2
4	$(m+n)r-\frac{(r+1)r}{2}$	$(n-1)(m+n)r - \frac{(n-1)(r+1)r}{2}$	$(n-1)(m+n)r - \frac{(n-1)(r+1)r}{2}$
5		$r^2(n-r)$	$r^2(n-r) + r$
6	?	?	?
7		nr^2	nr^2
8		mnr	mnr
total	$(m+n)r-\frac{(r+1)r}{2}$	$mn^2+nr^2+mnr +r^2(n-r) + +(n-1)r[(m+n)-\frac{r+1}{2}]$	$mn^2+nr^2+mnr +r^2(n-r)+r+ +(n-1)r[(m+n)-\frac{r+1}{2}]$
	+ no. of operations in inverting an r x r matrix		

3.2.2 Examining the above table we note that the case $r = \min\{m,n\} -1 = n - 1$ requires the greatest numbers of arithmetic operations. In fact, if $r = n - 1$ then

$$(m + \frac{n}{2})\ (n-1) \text{ divisions}$$

$$2mn^2 - mn + (n-1)^2\ [m + \frac{3}{2}n + 1] \text{ multiplications}$$

$$\text{and } 2mn^2 - mn + (n-1)^2\ [m + \frac{3}{2}n + 1] + (n-1) \text{ additions}$$

are needed, except for the inversion of the $(n-1) \times (n-1)$ matrix in Step 6.

3.2.3 If $r = \min\{m,n\}$, and this fact is recognized by the machine, then Steps 5-8 are not executed. When A is nonsingular, i.e. $m = n = r$, the method requires

$$(\frac{3}{2}n^2 - \frac{n}{2}) \text{ divisions}$$

$(\frac{5}{2}n^3 - 2n^2 + \frac{1}{2}n)$ multiplications

and $(\frac{5}{2}n^3 - 2n^2 + \frac{1}{2}n)$ additions.

Thus, in the nonsingular case the above method is less efficient than the ordinary matrix inversion methods. This is due to the formation of A^*A, which is not needed for inversion by elimination, if A is nonsingular.

3.3 Discussion

3.3.1 We give the results of some experiments with the present elimination method, on Bendix G-20, for Rutishauser's matrix $A_n = (a_{ij}) = (\binom{i+j}{i})$, $0 \leq i, j \leq n-1$, with condition number $\log P(A_n) \sim 4\, n \log 2$, e.g. [29]. As in [29], we have chosen the measure of the error as

$m = \max\limits_{i,j} |r_{ij}|$ where $R = (r_{ij}) = A_n X - I$ and X is the computed inverse of A_n.

With the smallest number ε recognized by the algorithm, $\varepsilon = 10^{-30}$, the results are:

	error	compilation time	computation time
$n = 8$	$m = 7.1 \times 10^{-3}$	69 sec.	18 sec.
$n = 28$	$m = 7.4 \times 10^{-1}$	69 sec.	176 sec.

The same results, approximately, were obtained for $\varepsilon = 10^{-15}$ and $\varepsilon = 10^{-50}$. Comparing these results with those of Newman and Todd

([29] p. 474), it is possible to evaluate, for the nonsingular case, the present elimination method. As expected, see 3.2.3 above, in the nonsingular case the present method does not compare favorably with the ordinary elimination methods.

3.3.2 As in other elimination methods, the correct rank recognition by the algorithm -- which in turn depends on the choice of ϵ -- is all important here. The rank recognition may be tested here by comparing zero rows in EA*AP and in EA*. Since the matric equation A*AX = A* is consistent, it follows that for any zero row in EA*AP the corresponding row in EA* must be zero. Such a test is clearly incomplete.

4. An Iteration Method for Computing A^+

In this section we discuss briefly the iteration method 2.2.12 and relate our computational experience with it.

4.1 The Convergence Rate

4.1.1 In [2] it was shown that

$$A^+ = \alpha \sum_{k=0}^{\infty} (I - \alpha A^*A)^k A^* \quad \text{for } 0 < \alpha < \frac{2}{\lambda_1}$$

This is the basis of the iteration method 2.2.12 whose n^{th} iteration gives:

$$X_n = \alpha \sum_{k=0}^{\infty} (I - \alpha A^*A)^k A^*$$

$$= X_{n-1} + \alpha A^*(I - AX_{n-1}) \qquad n = 1, 2, \ldots$$

where $\quad X_0 = \alpha A^*$ and $0 < \alpha < \dfrac{2}{\lambda_1}$

As $A^+ = \lim_{n\to\infty} X_n$, we have in the n^{th} iteration

$$\| A^+ - X_n \| = \| \alpha \sum_{k=n+1}^{\infty} (I - \alpha A^*A)^k A^* \| \leq$$

$$\leq \alpha \sum_{k=n+1}^{\infty} \| I - \alpha A^*A \|^k \| A^* \| =$$

$$= \frac{\alpha \| A^* \| \cdot \| I - \alpha A^*A \|^{n+1}}{1 - \| I - \alpha A^*A \|}$$

where $\| T \|$ is any multiplicative matrix norm (i.e. $\| ST \| \leq \| S \| \cdot \| T \|$).

4.1.2 A more rapidly convergent scheme was given in [2]: Let

$$Y_n = \alpha \{ I + (I - \alpha A^*A) \} \prod_{k=1}^{n-1} \{ I + (I - \alpha A^*A)^{2^k} \} A^*$$

where $0 < \alpha < \frac{2}{\lambda_1}$.

Then $Y_n \to A^+$ and

$$\| A^+ - Y_n \| \leq \frac{\alpha \| A^* \| \, \| I - \alpha A^*A \|^{2^n+1}}{1 - \| I - \alpha A^*A \|}$$

Still more rapidly convergent schemes, at the cost of more multiplications per iteration, can be given, e.g. using the results in [26].

4.2 The Optimal α

4.2.1 To the vector norm $\| x \| = (x, x)^{1/2} = (\sum_{i=1}^{n} | x_i |^2)^{1/2}$ there "corresponds" the matrix norm $\| T \|_0 = \sqrt{\lambda_1}$,where λ_1 is the greatest eigenvalue of T^*T. By "correspondence" here we mean that

$$\| T \|_0 = \sup_{x \neq 0} \frac{\| Tx \|}{\| x \|}$$

Since $\| I - \alpha A^*A \|$, as a function of the real parameter α is convex, and

$\| I - \alpha A^*A \|_0 = 1$ for $\alpha = 0$ and for $\alpha = \frac{2}{\lambda_1}$ it follows that $\| I - \alpha A^*A \|_0$ assumes a minimum for some α_0 in $(0, \frac{2}{\lambda_1})$.

Let
$$M = \sup_{x \in R(A^*)} \frac{(x, A^*Ax)}{(x,x)}$$

$$m = \inf_{x \in R(A^*)} \frac{(x, A^*Ax)}{(x,x)}$$

It is well known that $M = \lambda_1$ and $m > 0$. As in [1] the minimizing value α_0 must satisfy $1 - \alpha m = -(1 - \alpha M)$

therefore $\alpha_0 = \frac{2}{M + m}$

and $\| I - \alpha_0 A^*A \|_0 = \frac{M - m}{M + m}$

The error estimate given in 4.1.1 becomes, upon using the norm $\| \;\; \|_0$ and $\alpha = \alpha_0$:

$$\| A^+ - X_n \|_0 \leqq \frac{M}{m}^{1/2} \left(\frac{M - m}{M + m}\right)^{n+1}$$

<u>4.2.2</u> To avoid the computation of M and m, α_0 can be approximated by α_1: the value of α which minimizes $\| I - \alpha A^*A \|_1$, where $\| \;\; \|_1$ is the "sum of squares" norm:

$$\| T \|_1 = \left(\sum_{i,j} | t_{ij} |^2\right)^{1/2}.$$

Differentiating the function

$$\| I - \alpha A^*A \|_1 = \left(\sum_{i,j} \left(\delta_{ij} - \alpha \sum_k \bar{a}_{kj} \, a_{kj}\right)^2 \right)^{1/2}$$

with respect to α, we verify that:

$$\alpha_1 = \frac{\sum_{i,j} | a_{ij} |^2}{\sum_{i,j} \left| \sum_k \bar{a}_{ki} \, a_{kj} \right|^2} = \frac{\text{trace } A^*A}{\text{trace } (A^*A)^*A^*A}$$

4.3 Discussion

We have found the convergence of the iterative method 2.2.12 too slow to be of much practical value. E.g. for the Rutishauser matrix A_8, α_1 is 5×10^{-8}. It may be worthwhile to check the effect of some accelerated versions of 2.2.12; this will be reserved for a future study.

REFERENCES

[1] M. ALTMAN: Pacific Jour. of Math. 10, 1960, 1107-1113.

A. BEN-ISRAEL AND A. CHARNES:

[2] "Contributions to the Theory of Generalized Inverses" to be published in Jour. of the SIAM.

[3] "Generalized Inverses and the Bott-Duffin Network Analysis" to be published in Jour. Math. Anal. Appl.

[4] A. BEN-ISRAEL AND S. J. WERSAN: "An Elimination Method for Computing the Generalized Inverse of an Arbitrary Matrix" to be published in Jour. Assn. Comp. Mach.

[5] H. BIALY: Arch. Rat. Mech. Anal. 4, 1959, 166-176.

A. BJERHAMMAR:

[6] Trans. Royal Inst. Tech. Stockholm No. 49, 1951, 1-86.

[7] Bull. Geodesigue 1951, 188-220.

[8] Kungl. Tekn. Hogsk. Handl. No. 124, 1958.

[9] G. G. den BROEDER, JR. AND A. CHARNES: "Contributions to the Theory of a Generalized Inverse for Matrices" Purdue University, Lafayette, Ind., 1957.

[10] A. CHARNES AND A. BEN-ISRAEL: Notices Amer. Math. Soc. 10, 1, 1963, p. 135.

R. E. CLINE:

[11] "On the computation of the generalized inverse A^+ of an arbitrary matrix A, and the use of certain associated eigenvectors in solving the allocation problem" Purdue University, Lafayette, Ind., 1958.

[12] "Generalized Inverse Representation for Matrices partitioned as $A = [U, V]$" The University of Michigan, 1962.

[13] P. S. DWYER: Linear Computations, J. Wiley, New York, 1951.

[14] G. E. FORSYTHE: Bull. Amer. Math. Soc. 59, 1953, 299-329.

[15] GATE 20 Algebraic Compiler for the Bendix G-20, Carnegie Institute of Technology, The Computation Center, 1963.

T. N. E. GREVILLE:

[16] SIAM Review 1, 1959, 38-43.

[17] SIAM Review 2, 1960, 15-22.

M. R. HESTENES:

[18] Jour. of the SIAM 6, 1958, 51-90.

[19] Pacific Jour. of Math. 11, 1961, 225-245.

[20] Arch. Rat. Mech. Anal. 11, 1962, 138-194.

[21] A. S. HOUSEHOLDER: Principles of Numerical Analysis, McGraw-Hill, New York, 1953.

[22] Y. IJIRI: Goal Oriented Models for Accounting and Control, Ph.D. Thesis, Carnegie Institute of Technology, 1963.

R. E. KALMAN:

[23] Boletin Soc. Mat. Mexic. 1960, 102-119.

[24] "New Methods and Results in Linear Prediction and Filtering Theory" RIAS TR 61-1, Baltimore, Md., 1961.

[25] C. LANCZOS: Amer. Math. Monthly 65, 1958, 665-679.

[26] A. T. LONSETH: Bull. Amer. Math. Soc. 60, 1954, 415-430.

E. H. MOORE:

[27] Bull. Amer. Math. Soc. 26, 1920, 394-395.

[28] General Analysis, Part I, Mem. Amer. Philos. Soc. 1, 1953; especially pp. 197-209.

[29] N. NEWMAN AND J. TODD: Jour. of the SIAM 6, 1958, 466-476.

R. PENROSE:

[30] Proc. Camb. Philos. Soc. 51, 1955, 406-413.

[31] Proc. Camb. Philos. Soc. 52, 1956, 17-19.

[32] L. D. PYLE: "A Gradient Projection Method for Solving Programming Problems using the Generalized Inverse A^+ and the Eigenvectors of $I-A^+A$" Purdue University, Lafayette, Ind., 1958.

[33] W. QUADE: Ber. Math. Tagung Tubingen, 1946, 123-124.

C. R. RAO:

[34] Jour. Roy. Stat. Soc. B 24, 1962, 152-158.

[35] Sankhya 15, 1955, 253-280.

[36] R. RADO: Proc. Camb. Philos. Soc. 52, 1956, 600-601.

M. STOJAKOVIC:

[37] Bull. Soc. Roy. Sci. Liege. 21, 1952, 303-305.

[38] Bull. Soc. Math. Phys. Serbie 4, 1952, 1-2,9-23. See Math. Rev. 14, 1953, p. 443.

[39] C. R. Acad. Sci. Paris 236, 1953, 877-879.

[40] C. R. Acad. Sci. Paris 246, 1958, 1133-1135.

[41] H. S. WILF: Jour. of the SIAM 7, 1959, 149-151.

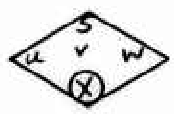
[42] S. ZACKS: "Generalized Least Squares Estimators for Randomized Fractional Replication Designs" Stat. and Appl. Math. Lab., Stanford University, Stanford, Calif., 1963.

[43] M. ZELEN: "The role of constraints in the theory of least squares" MRC Report No. 314, Math. Research Center U. S. Army, The University of Wisconsin, Madison, Wis., 1962.

APPENDIX 1: FLOW CHARTS

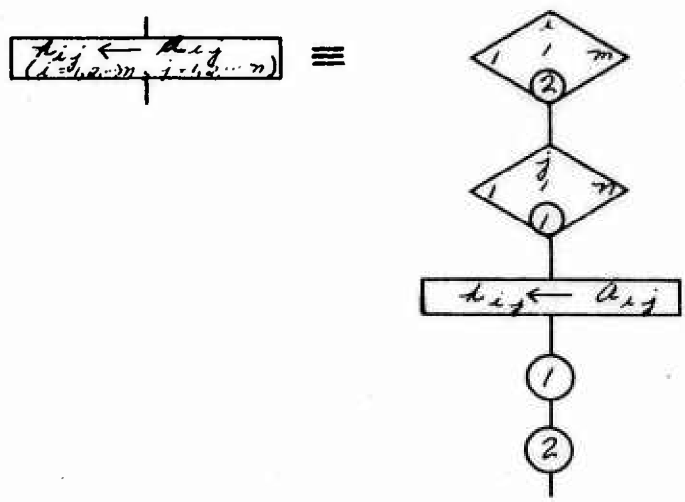
A:	Original m x n matrix whose g.i. is to be calculated
m:	Number of rows of A
n:	Number of columns of A
$\bar{A}$:	Reduced form of A by eliminating zero rows and zero columns of A, if any. If there is no zero rows or zero columns in A, $\bar{A} = A$
$\bar{m}$:	Number of rows of A (Of course $\bar{m} = m$ if $\bar{A} = A$)
$\bar{n}$:	Number of columns of A (Of course $\bar{n} = n$ if $\bar{A} = A$)
$\bar{\bar{A}}$:	Reduced and conditionally transposed form of A: i. e. $\bar{\bar{A}} = \bar{A}$ if $\bar{m} \geq \bar{n}$ and $\bar{\bar{A}} = \bar{A}^*$ if $\bar{m} < \bar{n}$
$\bar{\bar{m}}$:	Number of rows of $\bar{\bar{A}}$ ($\bar{\bar{m}} = \bar{m}$ if $\bar{\bar{A}} = \bar{A}$, $\bar{\bar{m}} = \bar{n}$ if $\bar{\bar{A}} = \bar{A}^*$)
$\bar{\bar{n}}$:	Number of columns of $\bar{\bar{A}}$ ($\bar{\bar{n}} = \bar{n}$ if $\bar{\bar{A}} = \bar{A}$, $\bar{\bar{n}} = \bar{m}$ if $\bar{\bar{A}} = \bar{A}^*$)
r:	Rank of $\bar{\bar{A}}$, which is equal to the rank of $\bar{\bar{A}}^*\bar{\bar{A}}$
ε:	Error allowance: i.e. in any matrix if an element has an absolute value which is less than ε, the machine regards the element as equal to zero
f_0	Flag 0: 0 if the problem is not the last one; 1 otherwise
f_1	Flag 1: 0 if $\bar{A} = A$ and 1 if $\bar{A} \neq A$
f_2	Flag 2: 0 if $\bar{\bar{A}} = \bar{A}$ and 1 if $\bar{\bar{A}} = \bar{A}^*$
f_3	Flag 3: 0 if $\bar{\bar{A}}^*\bar{\bar{A}}$ is singular and 1 if $\bar{\bar{A}}^*\bar{\bar{A}}$ is nonsingular
$p_1 \ldots p_m$	Flag for zero rows in A: $p_i = -i$ if i^{th} row is a zero row; $p_i = i$ otherwise
$q_1 \ldots q_n$	Flag for zero columns in A: $q_j = -j$ if j^{th} column is a zero column; $q_j = j$ otherwise.
$t_1 \ldots t_{\bar{\bar{n}}}$	Flag for row interchange in Stage 5. Initially $t_i = i$ $(i=1, 2 \ldots \bar{\bar{n}})$. If i^{th} row and j^{th} row are interchanged during the process in Stage 5, the value of t_i and t_j are interchanged, too; i.e., $t_i = j$, $t_j = i$.

a_{ij}	Element in i^{th} row j^{th} column of a
h_{ij}	Element in i^{th} row j^{th} column of a matrix stored in H register region h_{ij} may be designated by h_k where $k = (i-1) \times \bar{n} + j$
$\bar{n}$	Number of columns of a matrix stored in H register region
d	Temporary storage location
i, j, k, l	Indices



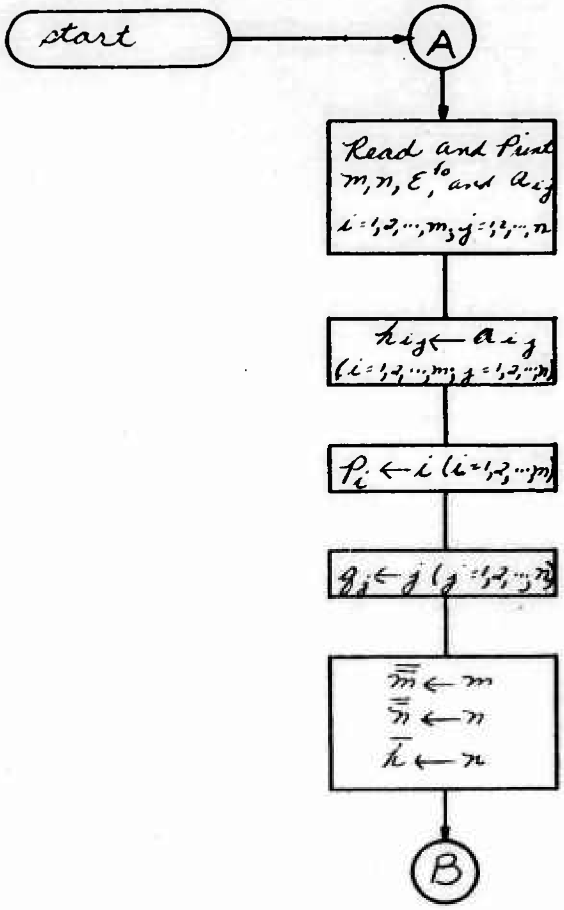
Iteration statement: Iterate the whole statements up to the point x in the diagram, setting the initial value of the index s equal to u and incrementing s by v at the end of each iteration, until s becomes equal to or greater than w. Then go to the statement immediately under the point x.

When there is no danger of confusion, an iteration statement is substituted by a statement showing the operation and the range for subscripts; e.g.



B-20

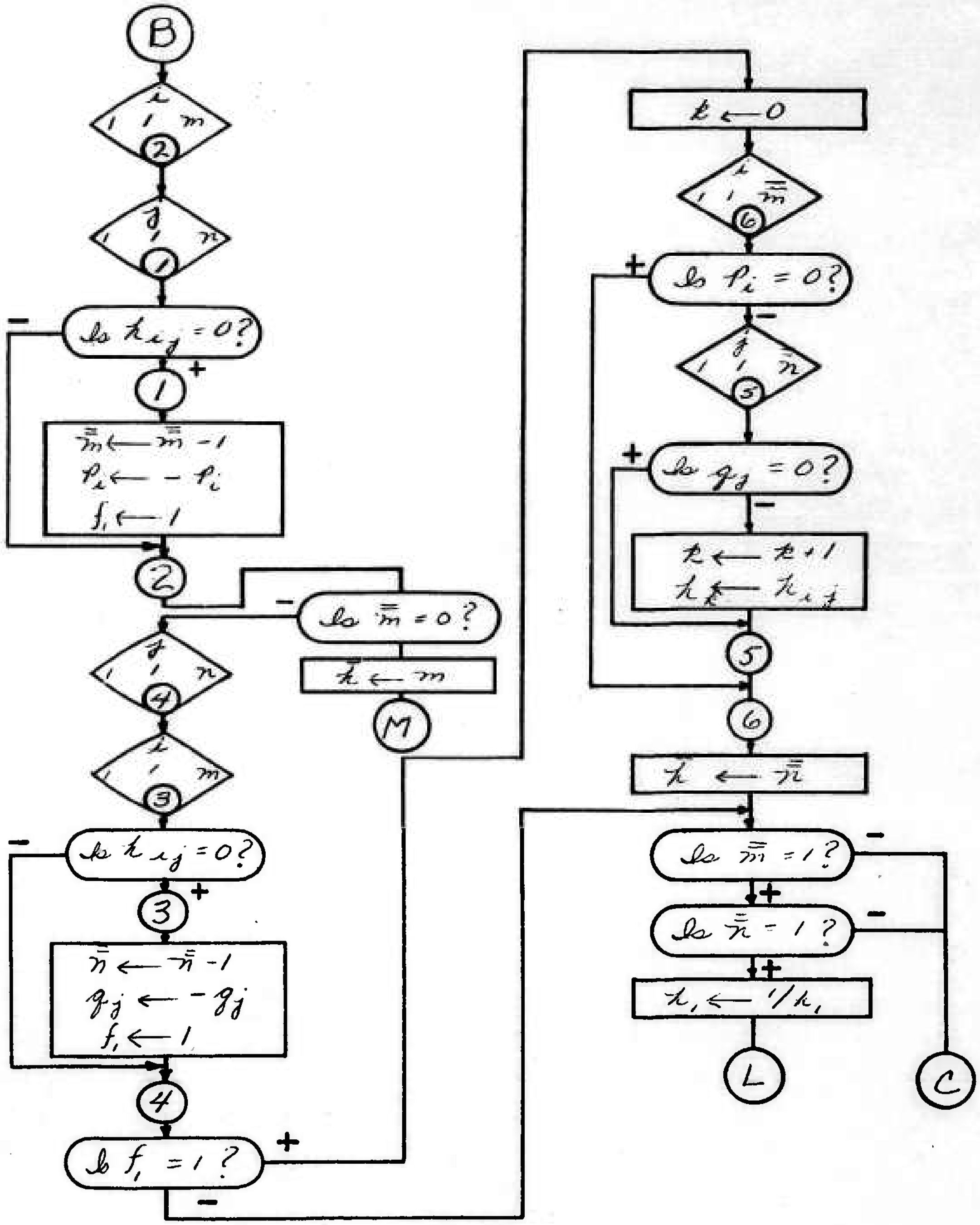
I. Initialization



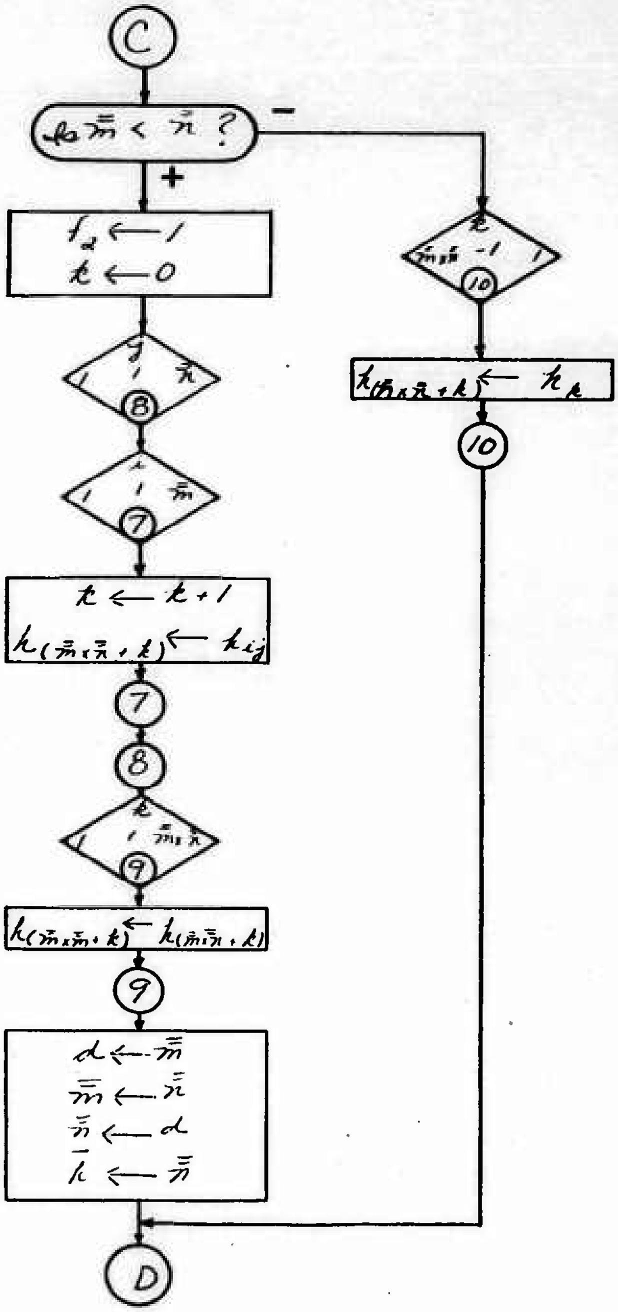
II. Eliminate Zero Rows and Columns

Find Zero Rows and Columns

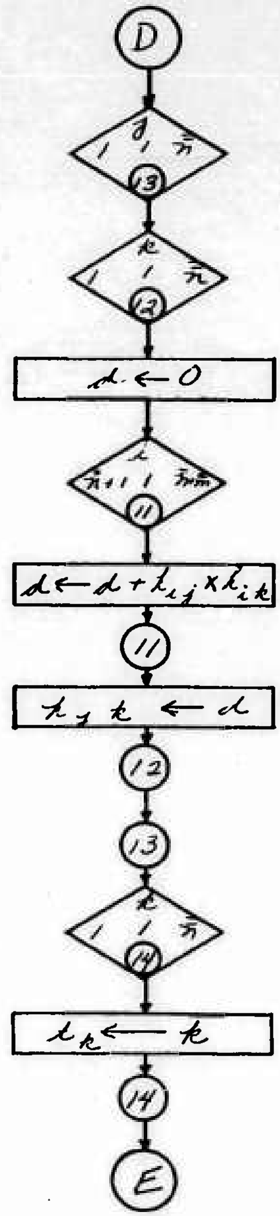
Eliminate Zero Rows and Columns



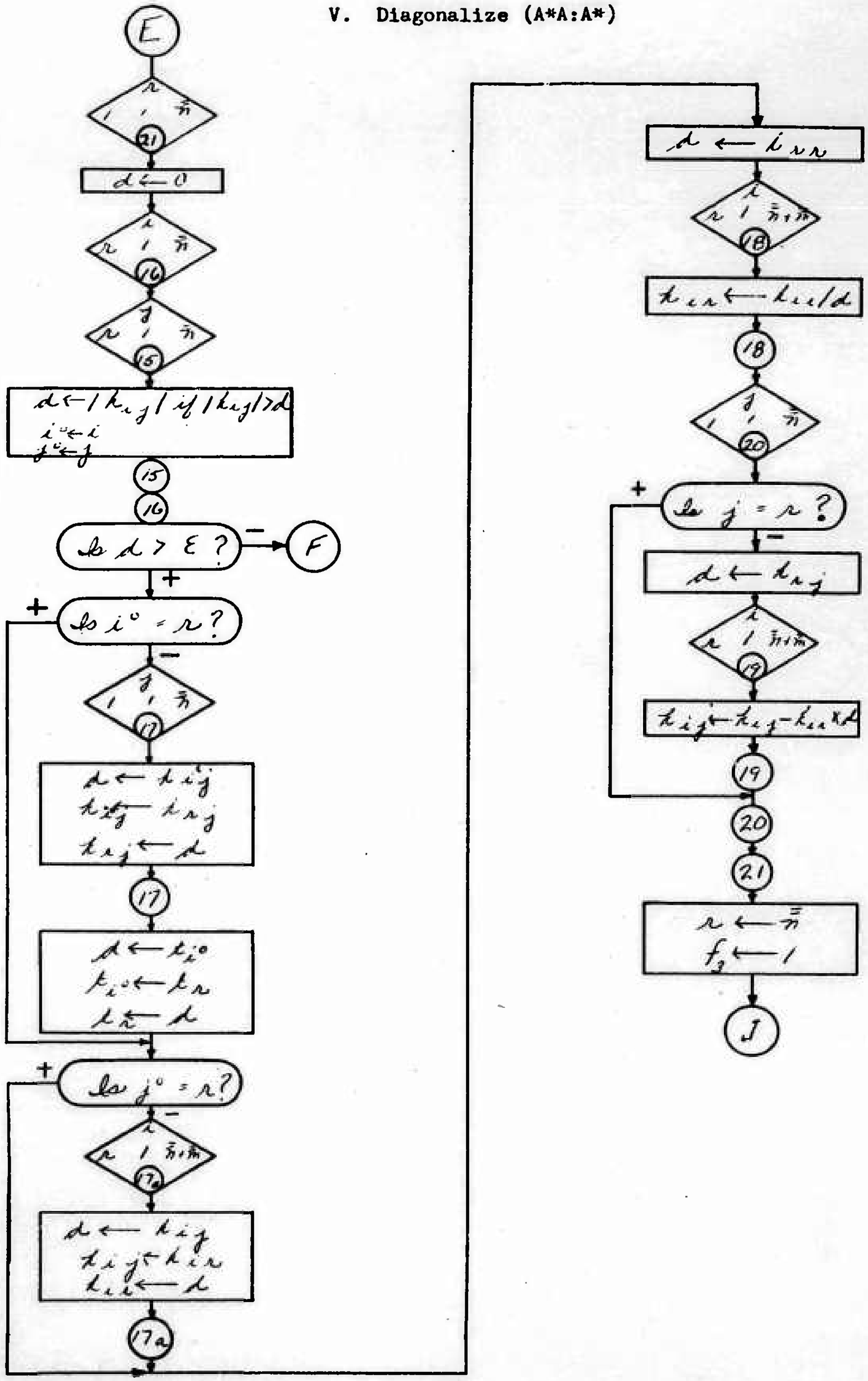
III. Transpose if $\bar{\bar{M}} < \bar{\bar{N}}$



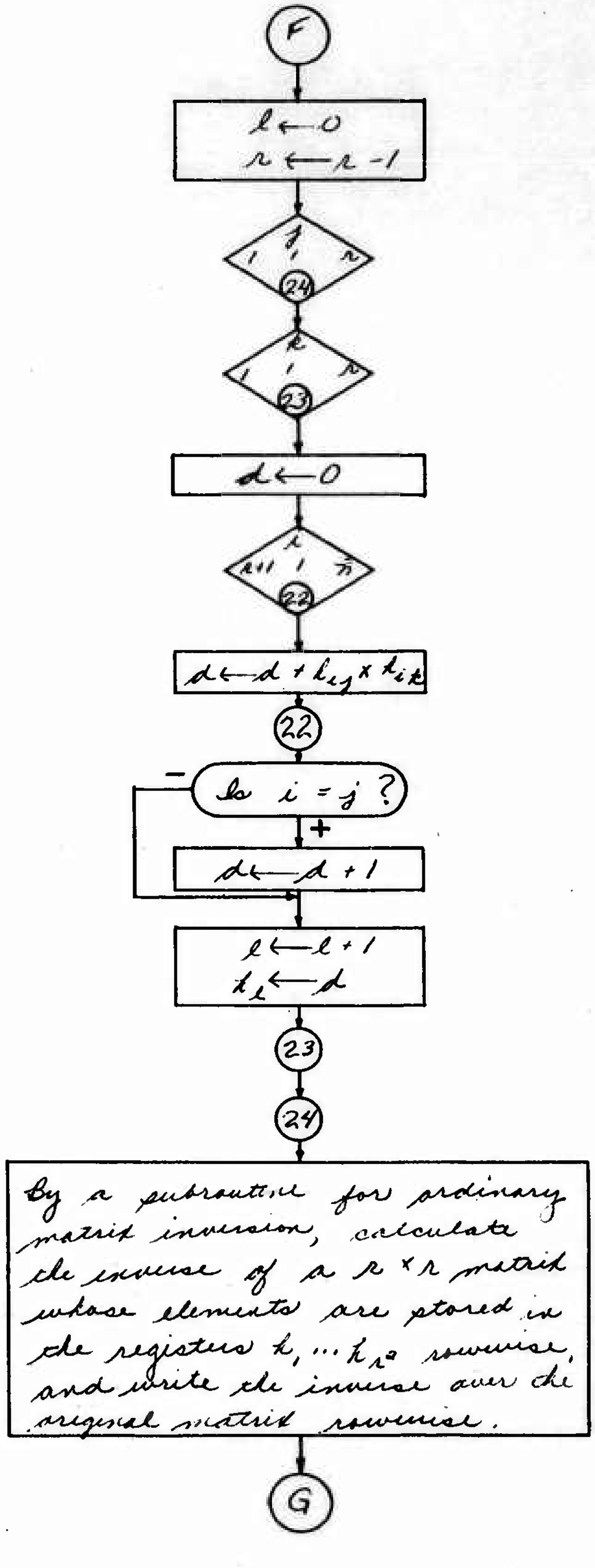
IV. Calculate $\bar{A}^*\bar{A}$



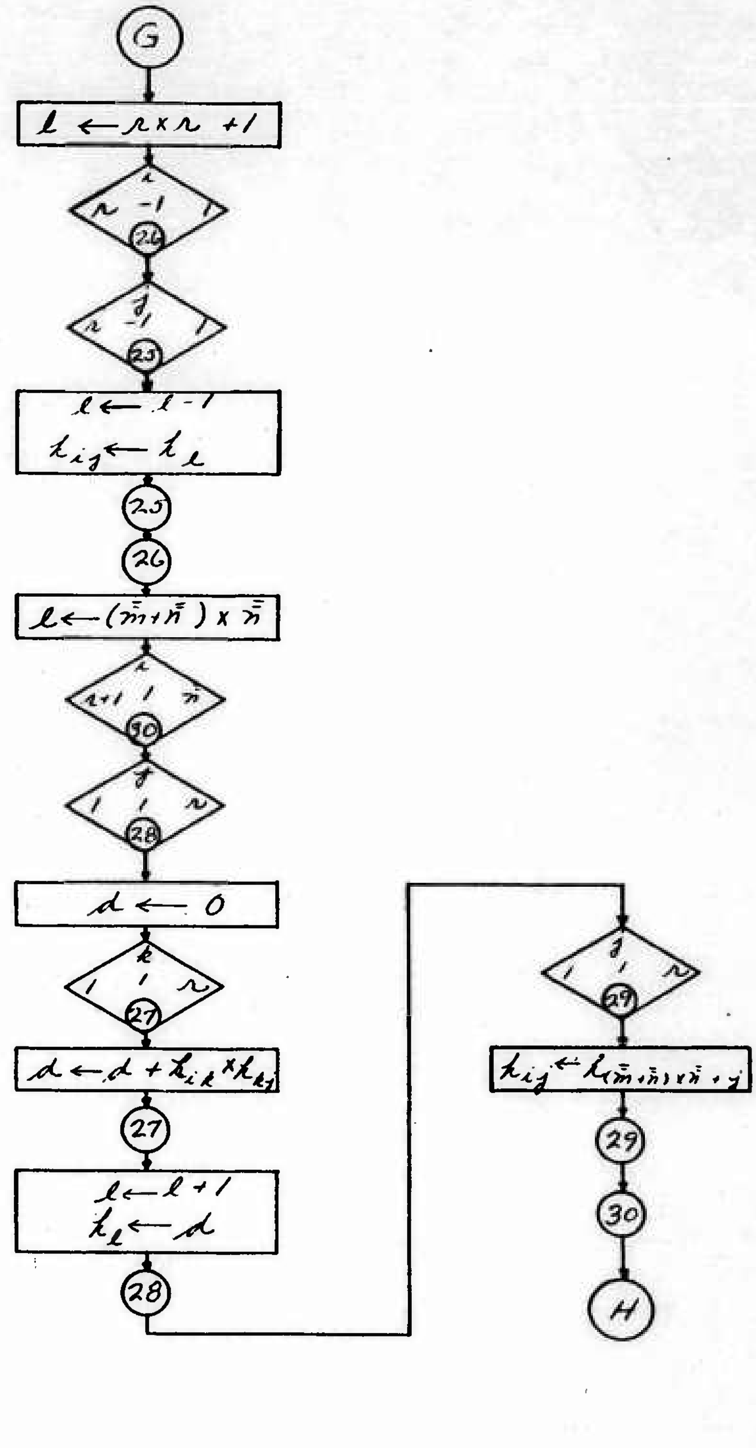
V. Diagonalize (A*A:A*)



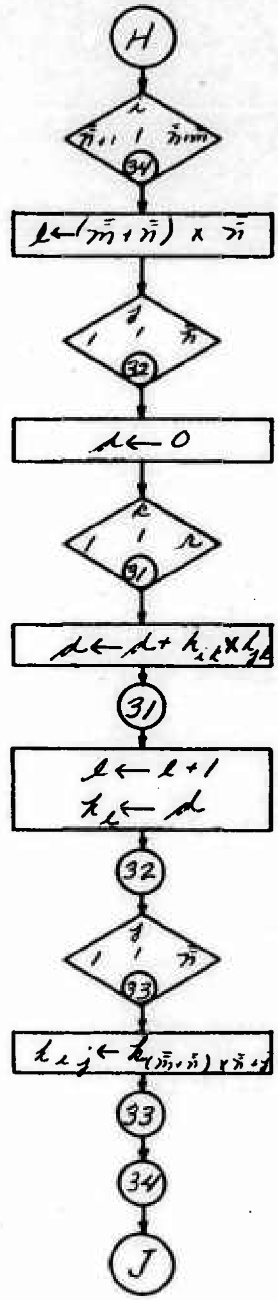
VI. Calculate $(I+\Delta\Delta^*)^{-1}$



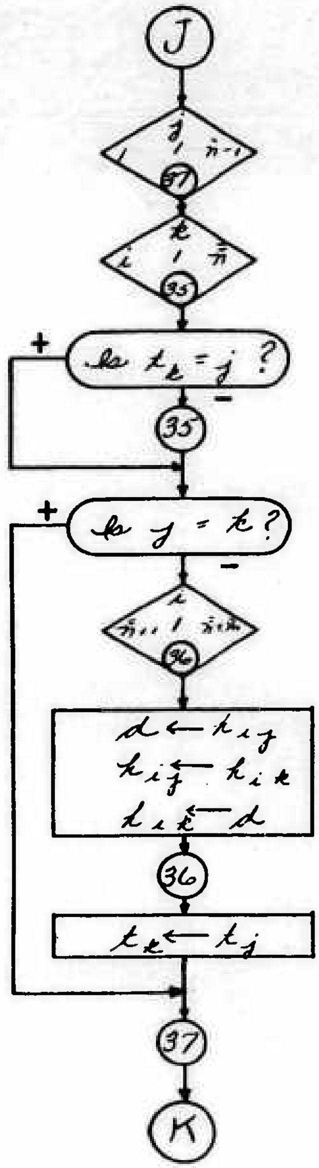
VII. Calculate $\binom{I}{\Delta^*}(I + \Delta\Delta^*)^{-1}$



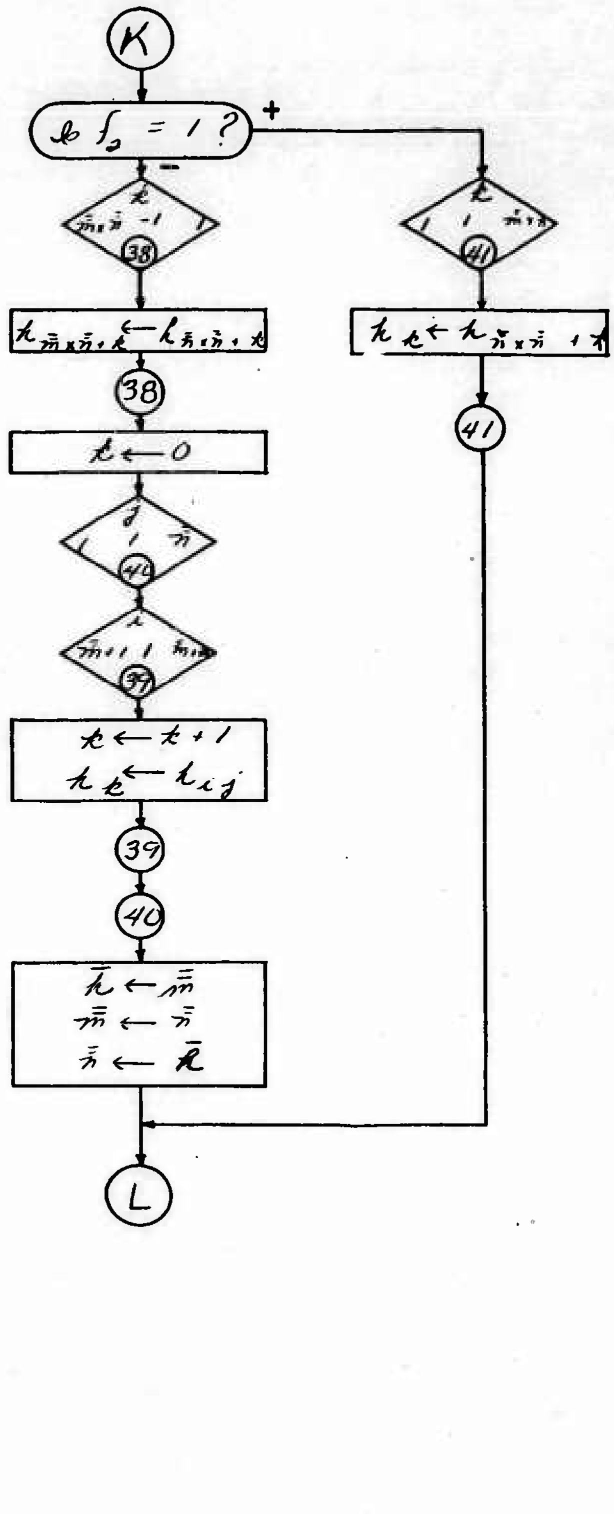
VIII. Calculate $(\frac{I}{\Delta}*)\ (I + \Delta\Delta*)^{-1}\ (EA*)$



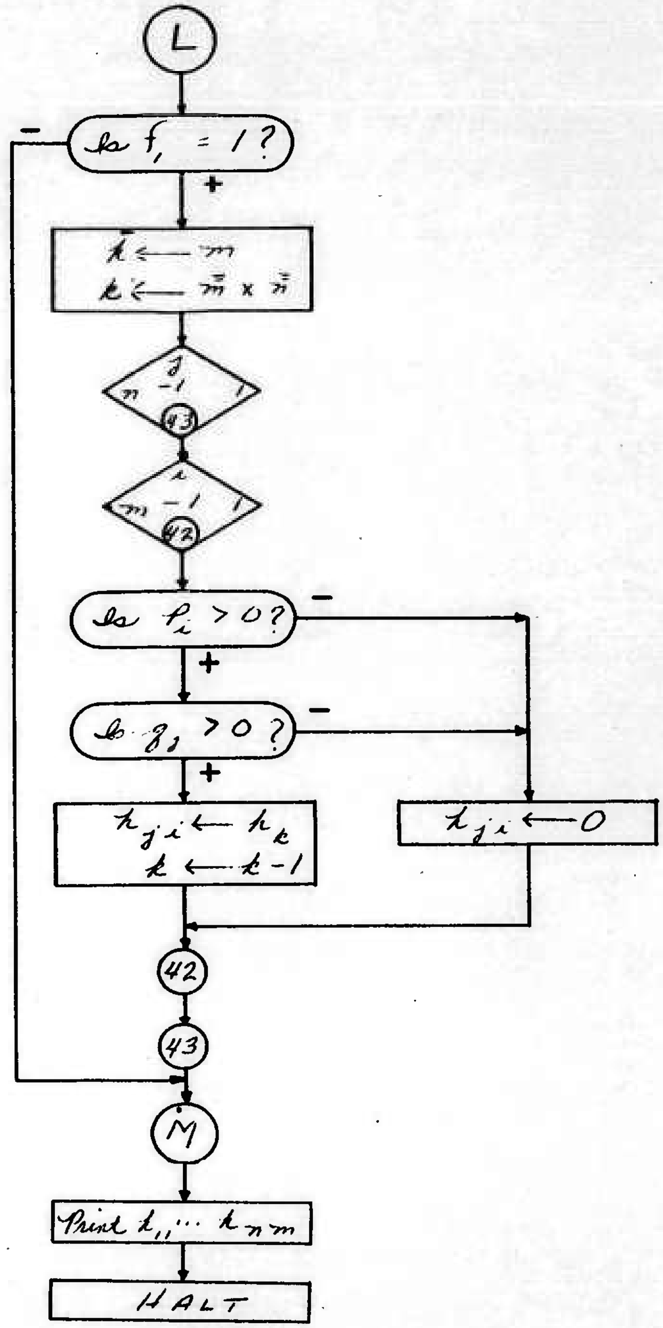
IX. Permute Columns if Necessary



X. Transpose if Necessary



XI. Insert Zero Rows or Columns if Any and Print A^+



APPENDIX 2:

```
      GATE PRØGRAM FØR GENERALIZED INVERSIØN ØF REAL MATRICES (GENINV)   M

      FØR EXPLANATIØN ØF SYMBØLS SEE [15]   M
      70 IS HIGHEST STATEMENT NUMBER
      DIMENSIØN C(1) D(1) G(5000,K8,1) H(10000,K14,1)   N
         I(100) K(30)


      INITIALIZATIØN   M
 1    K0=0 IF NØT END K1=M K2=N H0=E H1...H(M,N) C;READ
      K14 ← K2; K9 ← 0; K10 ← 0; K11 ← 0; K8 ← K2; K7 ← K1
      60, K20, 1, 1, K7*K8,
60    GK20 ← HK20
      TK1 TK2
      58, K20, 1, 1, K7,
      TK20
      58, K21, 1, 1, K8,
58    TH(K20, K21); TK21
      2, K20, 1, 1, K7,
 2    IK20 ← K20
      3, K20, 1, 1, K8,
 3    I(K7+K20) ← K20


      CHECK FØR ZERØ RØWS AND CØLUMNS   M
      5, K21, 1, 1, K7,
      4, K22, 1, 1, K8,
 4    GØ TØ 5 IF H(K21, K22) ≠ 0
      IK21 ← -IK21; K11 ← 1; K1 ← K1-1
 5    C
      GØ TØ 38; K14 ← K7 IF K1 = 0
      7, K22, 1, 1, K8,
      6, K21, 1, 1, K7,
 6    GØ TØ 7 IF H(K21, K22) ≠ 0
      I(K7+K22) ← -I(K7+K22); K11 ← 1; K2 ← K2 - 1
 7    C
      GØ TØ 12 IF K11 = 0
```

```
    ELIMINATE ZERØ RØWS ØR CØLUMNS   M
    K20 ← 0
    8, K21, 1, 1, K7,
    8, K22, 1, 1, K8,
8   HK20 ← H(K21, K22); K20 ← K20+1 IF IK21 > 0   N
       IF I(K7+K22) > 0
    K14 ← K2
12  K3 ← K1 + K2; K4 ← K1*K2; K6 ← K7 + K8
    GØ TØ 48; H1← 1/H1 IF K1 = 1 IF K2 = 1


    TRANSPØSE IF (M) < (N)   M
    K9 ← 1 IF K1 < K2
    GØ TØ 15 IF K9 = 0
    K20 ← K4
    13, K22, 1, 1, K2,
    13, K21, 1, 1, K1,
13  HK20 ← H(K21, K22); K20 ← K20+1
    14, K20, 1, 1, K4,
14  HK20 ← H(K20+K4)
    K2 ← K14; K1 ← K2; K14 ← K1
15  K5 ← K2*K2; K13 ← (K1+K2)*K2
    16, K20, K4, -1, 1,
16  H(K5+K20) ← HK20


    FØRM ((A))*X((A))   M
    19, K22, 1, 1, K2,
    19, K21, 1, 1, K2,
    DO ← 0
    18, K20, K2+1, 1, K3,
18  DO ← DO + H(K20, K21)*H(K20, K22)
19  H(K21, K22) ← DO
    20, K20, 1, 1, K2,
20  I(K6+K20) ← K20


    FIND AN ELEMENT WITH LARGEST ABSØLUTE VALUE   M
    9, K12, 1, 1, K2,
    DO ← AHO
    23, K18, K12, 1, K2,
    23, K19, K12, 1, K2,
23  K21 ← K18;K22 ← K19;DO ← AH(K18,K19)  IF AH(K18, K19) > DO
    GØ TØ 31; TK12 IF DO = HO
```

```
      RØW INTERCHANGE    M
24    GØ TØ 26 IF K21=K12
      25, K20, 1, 1, K2,
25    H(K12, K20) ← DO; H(K21, K20) ← H(K12, K20);    N
         DO ← H(K21, K20)
      I(K6+K12) ← IO; I(K6+K21) ← I(K6+K12);    N
         IO ← I(K6+K21)

      CØLUMN INTERCHANGE    M
26    GØ TØ 28 IF K22=K12
      27, K20, K12, 1, K3,
27    H(K20,K12) ← DO; H(K20, K22) ← H(K20, K12);    N
         DO ← H(K20, K22)

      DERIVATIØN    M
28    DO ← H(K12,K12)
      29, K20, K12+1, 1, K3,
29    H(K20,K12) ← H(K20,K12)/DO
      9, K22, 1, 1, K2,
      GØ TØ 9 IF K22=K12
      DO ← H(K12,K22)
      30, K21, K12+1, 1, K3,
30    H(K21,K22) ← H(K21, K22) - H(K21,K12)*DO
 9    C
      GØ TØ 11; K10 ← 1; K12 ← K2

      FØRM INVERSE ØF (I+(DELTA)(DELTA)*)    M
31    K12 ← K12-1; K20 ← 0
      32, K21, 1, 1, K12,
      32, K22, 1, 1, K12,
      DO ← 0
      33, K23, K12+1, 1, K2,
33    DO ← DO + H(K23,K21)*H(K23,K22)
      DO ← DO+1 IF K21=K22
32    HK20 ← DO; K20 ← K20+1
      MIVNC.  (LH1,K12)
      K20 ← K12*K12 + 1

      FØRM (I/(DELTA)*) X (ABØVE INVERSE)    M
      34, K21, K12, -1, 1,
      34, K22, K12, -1, 1,
34    H(K21,K22) ← HK20; K20 ← K20-1
      K20 ← K13
      10, K21, K12+1, 1, K2,
      36, K22, 1, 1, K12,
      DO ← 0
      35, K23, 1, 1, K12,
35    DO ← DO + H(K23, K22)*H(K21,K23)
36    HK20 ← DO; K20 ← K20+1
      K20 ← K13
      10, K22, 1, 1, K12,
10    H(K21, K22) ← HK20; K20 ← K20+1
```

```
    FØRM (EA*)* X (ABØVE MATRIX)*    M
    40, K21, K2+1, 1, K3,
    K20 ← K13
    39, K22, 1, 1, K2,
    DO ← 0
    37, K23, 1, 1, K12,
37  DO ← DO + H(K22,K23)*H(K21,K23)
39  HK20 ← DO; K20 ← K20+1
    40, K22, 1, 1, K2,
40  H(K21,K22) ← H(K13+K22)


    PERMUTING CØLUMNS    M
11  42, K21, 1, 1, K2-1,
    43, K22, K21, 1, K2,
43  GØ TØ 44 IF I(K6+K22)=K21
44  GØ TØ 42 IF K21 = K22
    45, K23, K2+1, 1, K3,
45  H(K23,K22) ← DO; H(K23,K21) ← H(K23,K22);    N
       DO ← H(K23,K21)
    I(K6+K22) ← I(K6+K21)
42  C


    TRANSPØSE IF NECESSARY    M
    GØ TØ 17 IF K9=0
    41, K20, 1, 1, K4,
41  HK20 ← H(K5+K20)
    GØ TØ 48
17  46, K20, K4, -1, 1,
46  H(K20+K4) ← H(K20+K5)
21  K20 ← 0
    47, K22, 1, 1, K2,
    47, K21, K1+1, 1, K1+K1,
47  HK20 ← H(K21,K22); K20 ← K20+1
    K2 ← K14; K1 ← K2; K14 ← K1


    INSERT ZERØS IF NECESSARY    M
48  GØ TØ 38 IF K11=0
    K20 ← K4; K14 ← K7
    49, K22, K8, -1, 1,
    49, K21, K7, -1, 1,
    GØ TØ 49; K20 ← K20-1; H(K22,K21) ← HK20    N
       IF I(K7+K22) > 0 IF IK21 > 0
    H(K22,K21) ← 0
49  C
```

```
     PRINT (A+)   M
38   C
     22, K20, 1, 1, K8,
     TK20
     22, K21, 1, 1, K7,
22   TH(K20,K21); TK21
     GØ TØ 1 IF K0=0
     HALT
     PRØGRAM END
```

APPENDIX C

A Computational Algorithm for Generating All the Extreme Points of a Convex Set Mapped into a Two Dimensional Space

1. Introduction

The algorithm described here was developed as an extension of a dual evaluator analysis of a linear programming problem. The algorithm is designed to supply enough data from which the maximum value $\hat{v}$ and the minimum value $\check{v}$ of the functional v of a linear programming problem are determined for any given value of a stipulation level w of a given constraint in the linear programming problem.

Since the loci of $\hat{v}$ and $\check{v}$, as w changes, form a convex set in a two-dimensional space $\mathcal{X}_2$, the algorithm generates all the extreme points of the convex set, which may be considered as a convex set in an n-dimensional space $\mathcal{X}_n$ mapped into $\mathcal{X}_2$.

2. Problem

Let B (mxn), h (mx1), a (1xn), and c (1xn) be a given matrix and given vectors and x (nx1) be a vector of variables in $\mathcal{X}_n$. Let S be a non-empty bounded convex set in $\mathcal{X}_n$ defined by the system of linear inequalities,

$$Bx \leq h$$
$$x \geq 0.$$

Also let $\hat{S}$ be a non-empty bounded convex set obtained by mapping S into $\mathcal{X}_2$ whose axes are w and v where

$$w \equiv cx$$

and

$$v \equiv ax.$$

Our problem is to develop an algorithm which generate all the extreme points of the mapped convex set S.

3. Algorithm

1. Notation

(LP.0): The linear programming problem,

$$\begin{aligned} \text{Minimize} \quad & w = cx \\ \text{subject to} \quad & Bx \leq h \\ & x \geq 0. \end{aligned}$$

(LP.1): The linear programming problem,

$$\begin{aligned} \text{Maximize} \quad & v = ax - My \\ \text{subject to} \quad & Bx \leq h \\ & cx + y = w_1 \\ & x, y \geq 0. \end{aligned}$$

M: An unspecified but dominantly large positive number

w_1: The optimum value of the functional of (LP.0)

v_1: The optimum value of the functional of (LP.1)

T^k: A simplex tableau at the beginning of the step (3a.k) in the iteration in the algorithm. ($k=1,2,\ldots$). In particular, T^0 refers to an initial tableau for (LP.1), and T^1 an optimum tableau to (LP.1).

P_0^k: The stipulation vector in T^k

P_y^k: The vector for the artificial variable y in T^k

P_j^k: The vector for the structural and slack variables in T^k ($j=1,2,\ldots, n+m$).

x_{i0}^k, x_{iy}^k, x_{ij}^k: The i^{th} element in P_0^k, P_y^k, and P_j^k, respectively ($i=1,2,\ldots, m+1$).

x_{0j}^k: The dual evaluator, z_j-c_j, in P_j ($j=1,2,\ldots, n+m$)

x_{0y}^k: The dual evaluator, z_j-c_j, in P_y, excluding M.

Δw_k: The minimum value of $x_{i0}^k/|x_{iy}^k|$ among all i whose $x_{iy}^k < 0$ and $x_{i0}^k > 0$, ($i=1,2,\ldots, m+1$).

$\hat{i}^k$: The index for the row from which Δw_k was derived

$\hat{j}^k$: The index for the column whose $x_{0j}^k/|x_{\hat{i}^k j}^k|$ is minimum among all j whose $x_{\hat{i}^k j}^k < 0$, ($j=1,2,\ldots n+m$).

$\hat{P}_0^k$: $= P_0^k + \Delta w_k P_y^k$, (M in the dual evaluator for P_y must be excluded in the calculation of $\hat{P}_0^k$.)

w_{k+1}, v_{k+1}: $w_{k+1} = w_k + \Delta w_k$; $v_{k+1} = v_k + \Delta w_k x_{0y}^k$.

2. <u>Algorithm</u> 3. Algorithm

(1) Derive w_1 from (LP.0).

(2) Formulate an initial simplex tableau, T^0, for (LP.1), derive an optimum tableau, T^1, and record v_1.

(3) Repeat (3a.k)-(3c.k) for k=1,2,... until the artificial vector y is forced into the basis in (3c.k).

(3a.k): Calculate Δw_k and w_{k+1}, and record w_{k+1}.

(3b.k): Calculate P_0 and record v_{k+1}.

(3c.k): Pivot on the element in the $\hat{i}^k$ th row and the $\hat{j}^k$ th column and obtain T^{k+1}.

(4) Repeat (2)-(3) after changing the functional in (LP.1) to "Maximize $-v=-ax-My$."

4. An Example

We shall show the process of the algorithm by the following example.

Problem: Generate all the extreme points of the convex set S obtained by mapping the convex set S[1] into a two dimensional space whose axes are w and v, where S is defined by the system of linear inequalities,

$$\begin{aligned} 3x_1 + 2x_2 - x_3 &\le 6 \\ 3x_1 + 2x_2 + 4x_3 &\le 16 \\ 3x_1 \qquad - 4x_3 &\le 3 \\ 9/4x_1 + 4x_2 + 3x_3 &\le 17 \\ x_1 + 2x_2 + x_3 &\le 10 \\ x_1, x_2, x_3 &\ge 0 \end{aligned} \tag{C.1}$$

and

$$\begin{aligned} w &= x_1 + x_2 + x_3 \\ v &= x_1 + 2x_2 + 3x_3. \end{aligned} \tag{C.2}$$

---

[1] This convex set was taken from Balinski, 1961.

We first solve $(LP.\hat{0})$ using the above data, i.e.

$$\begin{aligned} \text{Minimize } w &= x_1 + x_2 + x_3 \\ \text{subject to } 3x_1 + 2x_2 - x_3 &\le 6 \\ 3x_1 + 2x_2 + 4x_3 &\le 16 \\ 3x_1 \qquad - 4x_3 &\le 3 \\ 9/4x_1 + 4x_2 + 3x_3 &\le 17 \\ x_1 + 2x_2 + x_3 &\le 10 \\ x_1, x_2, x_3 &\ge 0 \end{aligned} \tag{LP.0}$$

The minimum value of w, denoted by w_1, to the above problem is equal to zero. (The step (1) has been completed.)

Then, using this value of w_1, we formulate (LP.$\hat{1}$) as follows:

$$
(\text{LP.}\hat{1}) \qquad
\begin{array}{lrl}
\text{Maximize} & v = & x_1 + 2x_2 + 3x_3 - My \\
\text{subject to} & & 3x_1 + 2x_2 - x_3 \leq 6 \\
& & 3x_1 + 2x_2 + 4x_3 \leq 16 \\
& & 3x_1 \quad\quad - 4x_3 \leq 3 \\
& & 9/4x_1 + 4x_2 + 3x_3 \leq 17 \\
& & x_1 + 2x_2 + x_3 \leq 10 \\
& & x_1 + x_2 + x_3 + y = 0 \\
& & x_1, x_2, x_3 \geq 0
\end{array}
$$

We shall show the step (3) of the algorithm by a set of simplex tableaus T^0, T^1, ..., using a Tucker's (1958) condensed form of simplex tableaus eliminating all the basis vectors from the tableaus.

T^0: Following the ordinary simplex iterations, we obtain an optimum tableau T^1. The element marked by * is the pivot to obtain T^1

	c_j		1	2	3
c_i	B^0	P_0^0	P_1^0	P_2^0	P_3^0
	P_4	6	3	2	−1
	P_5	16	3	2	4
	P_6	3	3	0	−4
	P_7	17	9/4	4	3
	P_8	10	1	2	1
−M	P_y	0	1	1	1*
	$z_j - c_j$	0	$-1 \atop -M$	$-2 \atop -M$	$-3 \atop -M$

		1	2	$-M$	
B^1	P_0^1	P_1^1	P_2^1	P_y^1	$\hat{P}_0^1$
P_4	6	4	3	1	10
P_5	16	-1	-2*	-4	0
P_6	3	7	4	4	19
P_7	17	-3/4	1	-3	5
P_8	10	0	1	-1	6
3 P_3	0	1	1	1	4
$z_j - c_j$	0	2	1	$3+M$	12

T^1: From this optimum tableau, we obtain $v_1 = 0$. Thus $(w_1, v_1) = (0, 0)$ is an extreme point of $\hat{S}$. (The step (2) has been completed.)

(3a.1). Search for all the negative element in P_y (-4, -3, and -1). Divide the corresponding elements in P_0 by the absolute value of the negative elements just found. (16/4, 17/3, and 10/1). Take the minimum value of the quotients as Δw_1. ($\Delta w_1 = 16/4 = 4$). Calculate $w_2 = w_1 + \Delta w_1 = 0 + 4 = 4$.

(3b.1) Calcultae $\hat{P}_0^1 = P_0^1 + 4P_y^1$. The result is given in the last column of T^1. v_2 is 12. Thus, an extreme point $(w_2, v_2) = (4, 12)$ was obtained.

(3c.1) Since Δw_1 was derived from P_5 row, we search negative elements in P_5 row other than the element in P_y column. (-1 and -2). Divide the corresponding dual evaluators by the absolute value of the negative numbers just found. (2/1 and 1/2). Take the negative element whose quotient is minimum as the pivot to obtain T^2. (The element thus chosen is marked by * in T^1.) In this pivoting, we discard P_0^1 and take $\hat{P}_0^1$ as the stipulation level. Since the element in P_0 which corresponds to the pivotal element is always equal to zero, $\hat{P}_0^k = P_0^{k+1}$.

			1		−M	
	B^2	P_0^2	P_1^2	P_5^2	P_y^2	$\hat{P}_0^2$
	P_4	10	5/2	3/2	−5	5
2	P_2	0	1/2	−1/2	2	2
	P_6	19	5	2	−4	15
	P_7	5	−5/4*	1/2	−5	0
	P_8	6	−1/2	1/2	−3	3
3	P_3	4	1/2	1/2	−1	3
	$z_j - c_j$	12	3/2	1/2	1+M	13

T^2: (3a.2) $\Delta w_2 = 1$ taken from P_7 row. $w_3 = w_2 + \Delta w_2 = 5$.

(3b.2) $\hat{P}_0^2 = P_0^2 + P_y^2$ as shown in the last column. $v_3 = 13$, obtaining an extreme point $(w_3, v_3) = (5, 13)$.

(3c.2) Since the element in P_1 column is the only negative element in P_7 row from which Δw_2 was taken, (disregarding the element in P_y column and in P_7 row), this is taken as the pivot to obtain T^3.

					−M	
	B^3	P_0^3	P_7^3	P_5^3	P_y^3	$\hat{P}_0^3$
	P_4	5	2	5/2	−15	0
2	P_2	2	2/5	−3/10	0	2
	P_6	15	4	4	−24	7
1	P_1	0	−4/5	−2/5	4	4/3
	P_8	3	−2/5	3/10	−1	8/3
3	P_3	3	2/5	7/10	−3	2
	$z_j - c_j$	13	6/5	11/10	−5+M	34/3

T^3: (3a.a) $\Delta w_3 = 1/3$ taken from P_4 row. $w_4 = w_3 + \Delta w_3 = \frac{16}{3}$.

(3b.3) $\hat{P}_0^3 = P_0^3 + 1/3 P_y^3$ as shown in the last column.

$v_4 = 34/3$, obtaining an extreme point $(w_4, v_4) = (16/3, 34/3)$.

(3c.3) Since the only negative element in P_4 row from which Δw_3 was taken is in P_y column, P_y vector must be forced into the basis in T^4, indicating that the step (3) has been completed.

We then change the functional in (LP.1) to "Maximize $-v = -x_1 - 2x_2 - 3x_3 - My$." and repeat the step (2) and (3) as follows:

	B	P_0	P_1 (−1)	P_2 (−2)	P_3 (−3)
	P_4	6	3	2	−1
	P_5	16	3	2	4
	P_6	3	3	0	−4
	P_7	17	9/4	4	3
	P_8	10	1	2	1
−M	P_y	0	1^*	1	1
	$z_j - c_j$	0	1 −M	2 −M	3 −M

T^0: By the ordinary simplex iterations, we obtain an optimum tableau T^1. The element with * is the pivot to obtain T^1.

	B'	P_0'	P_y' (−M)	P_2' (−2)	P_3' (−3)	$\hat{P}_0'$
	P_4	6	−3	−1	−4	3
	P_5	16	−3	−1	1	13
	P_6	3	−3	−3	-7^*	0
	P_7	17	−9/4	7/4	3/4	59/4
	P_8	10	−1	1	0	9
−1	P_1	0	1	1	1	1
	$z_j - c_j$	0	−1 +M	1	2	−1

T^1: $v_1 = 0$, obtaining an extreme point $(w_1, v_1) = (0, 0)$. This extreme point has already been obtained in T_1.

(3a.1) $\Delta w_1 = 1$ taken from P_6 row. $w_2 = 1$

(3b.1) $\hat{P}^1_0 = P^1_0 + P^1_y$. $-v_1 = -1$, obtaing an extreme point $(w_1, v_1) = (1, 1)$.

(3c.1) The element with * is the pivot.

	B^2	P_0^2	P_y^2 ($-M$)	P_2^2 (-2)	P_6^2	$\hat{P}_0^2$
	P_4	3	$-9/7$	$4/7$	$-4/7$*	0
	P_5	13	$-24/7$	$-10/7$	$1/7$	5
-3	P_3	0	$3/7$	$3/7$	$-1/7$	1
	P_7	$59/4$	$-18/7$	$10/7$	$3/28$	$35/4$
	P_8	9	-1	1	0	$20/3$
-1	P_1	1	$4/7$	$4/7$	$1/7$	$7/3$
	$z_j - c_j$	-1	$-13/7 + M$	$1/7$	$2/7$	$-16/3$

T^2: $\Delta w_2 = 7/3$ from P_4.

$(w_3, v_3) = (\frac{10}{3}, \frac{16}{3})$.

	B^3	P_0^3	P_y^3 ($-M$)	P_2^3 (-2)	P_4^3	$\hat{P}_0^3$
	P_6	0	$9/4$	$-5/4$	$-7/4$	3
	P_5	5	$-15/4$	$-5/4$*	$1/4$	0
-3	P_3	1	$3/4$	$1/4$	$-1/4$	2
	P_7	$35/4$	$-45/16$	$25/16$	$3/16$	5
	P_8	$20/3$	-1	1	0	$16/3$
-1	P_1	$7/3$	$1/4$	$3/4$	$1/4$	$8/3$
	$z_j - c_j$	$-16/3$	$-10/4 + M$	$1/2$	$1/2$	$-26/3$

T^3: $\Delta w_3 = 4/3$ from P_5.

$(w_4, v_4) = (\frac{14}{3}, \frac{26}{3})$.

	B^4	P_0^4	P_y^4 ($-M$)	P_5^4	P_4^4	$\hat{P}_0^4$
	P_6	3	6	-1	-2	7
-2	P_2	0	3	$-4/5$	$-1/5$	2
-3	P_3	2	0	$1/5$	$-1/5$	2
	P_7	5	$-15/2$	$5/4$	$1/2$	0
	P_8	$16/3$	-4	$4/5$	$1/5$	$8/3$
-1	P_1	$8/3$	-2	$3/5$	$2/5$	$4/3$
	$z_j - c_j$	$-31/3$	$-4 + M$	$2/5$	$3/5$	$-34/3$

T^4: $\Delta w_4 = 2/3$ from P_7.
$(w_5, v_5) = (16/3, 34/3)$.
Since the only negative element in P_7 row is in P_y column, P_y must be forced into the basis in T^5, which indicates that the iteration has been completed.

The set of all the extreme points of the original convex set given by (C.1) is (0, 0, 0), (0, 3, 0), (0, 7/2, 1), (0, 2, 3), (0, 0, 4), (8/3, 0, 2), (1, 3/2, 0), (7/3, 0, 1), (4/3, 2, 2), and (1, 0, 0),[1] which are transformed into the two dimensional

[1] Balinski, 1961.

space by the transformation given in (C.2) as follows: (0, 0), (3, 6), (9/2, 10), (5, 13), (4, 12), (14/3, 26/3), (5/2, 4); (10/3, 16/3), (16/3, 34/3), and (1, 1). These points are plotted on the following Figure C-1. Note that not all the extreme points in the original convex set are the extreme points of the mapped convex set. Some are in the interior. But all the extreme points in the mapped convex set must come from extreme points in the original convex set.[1] The algorithm generates only the extreme

[1] See Theorem 5, p.237, in Charnes and Cooper, 1961.

points in the mapped convex set.

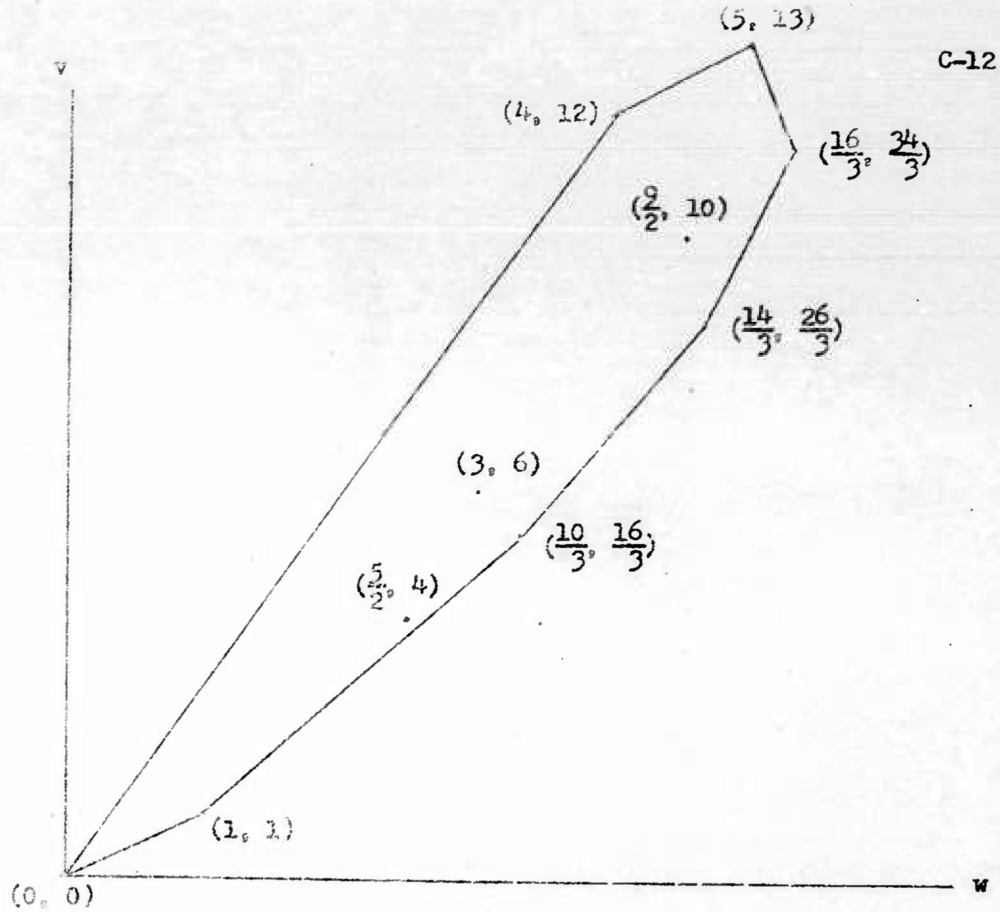


Figure C-1

Mapped Convex Set

5. Discussion

There have been already a few methods devised for generating all the extreme points of an n-dimensional convex set.[1]

[1] See discussions of the double description method in Motzkin, Raiffa, Thompson, and Thrall, 1953. See the use of a Tarry procedure in Charnes, 1952, and the spiral method in Charnes and Cooper, 1961, Ch. IX. See finally Balinski, 1961.

The algorithm described here, however, possesses advantages for managerial uses of the kind discussed in Chapter IV and VI. It requires only a minor change in a computer program for the simplex method.

The algorithm starts with extreme points in the mapped convex set which have the minimum value of $w = w_1$. There are at most two such extreme points, between which the algroithm takes the one with higher value of v. It then checks how much w may be increased without changing the basis, and derive an optimum tableau for (LP.1) in which w_1 is substituted by w_k ($k = 2, 3, \ldots$) via the successive bases. The step (3b.k) is a short cut to obtain such optimum tableaus without starting the simplex iterations afresh.[1]

[1] See Charnes and Cooper, 1961, p. 170 ff.

This is continued until a stage is reached at which P_y, the artificial vector, will be forced into the basis. This means that

we must leave the convex set in order to further increase w, and, therefore, we stop the iteration at that stage.

This completes the search for all the extreme points in the upper boundary of the mapped convex set. By changing the functional to "Minimize $v = ax + My$" or equivalently "Maximize $-v = -ax - My$," the algorithm generates all the extreme points in the lower boundary.

The only possibility of generating an extreme point twice is that the initial and final extreme points generated during the step (4) may have already been generated in the earlier step (3) due to the fact that the same extreme point may be at the intersection of the upper and the lower boundary of the mapped convex set.

When the original convex set S is empty, the mapped convex set $\hat{S}$ is also empty. When the original convex set is unbounded, they may be closed by the so-called "regularization method"[1] without

[1] See Charnes and Cooper, 1961.

affecting the nature of the problem and the solutions, and so there is no trouble here either and the same remark applies via these regularization procedures when the original problem has no solution.

The possibility of extending the algorithm to cases where the mapped convex set has more than two dimensions remains to be explored, since this is of possible interest when multiple "indicators" and their related "divergence coefficients" are to be accomodated.

BIBLIOGRAPHY

American Institute of Accountants. Changing Concepts of Business Income (a report of the Study Group on Business Income). New York: The Macmillan Co., 1952.

Anthony, Robert N. Management Accounting: Text and Cases (Revised Edition). Homewood, Illinois: Richard D. Irwin, Inc., 1960.

Balinski, M.L. "An Algorithm for Finding All Vertices of Convex Polyhedral Sets," Journal of the Society for Industrial and Applied Mathematics, IX, No.1, March, 1961, pp. 72-88.

Baridon, Felix E. "Profit and Loss Budget by Volume," NAA Bulletin, Sec. 1, November, 1960, pp. 83-89.

Baxter, W.T. and S. Davidson (eds.). Studies in Accounting Theory, 2nd ed. Homewood, Illinois: Richard D. Irwin, Inc., 1962.

Bedford, Norton M., Charles H. Griffin, and Thomas H. Williams. "Emerging Tools for Managerial Accounting," NAA Bulletin, XLIII, October, 1961, pp. 31-37.

_______, _______, and ______. "Emerging Role of Mathematical Methodology in Accounting," NAA Bulletin, XLIII, June, 1962, pp. 33-38.

Ben-Israel, A. On Some Problems of Mathematical Programming. Ph.D. Thesis. Northwestern University, June, 1962.

_______ and A. Charnes. "Generalized Inverses and the Bott-Duffin Network Analysis" to be published in Journal of Mathematical Analysis and Applications.

_______ and _______. "Contributions to the Theory of Generalized Inverses" to be published in Journal of the Society for Industrial and Applied Mathematics, 1963.

_______ and Y. Ijiri. "A Report on the Machine Computation of the Generalized Inverse of an Arbitrary Matrix," ONR Research Memorandum No.110. Pittsburgh: Carnegie Institute of Technology, March, 1963.

_______ and S.J. Wersan. "An Elimination Method for Computing the Generalized Inverse of an Arbitrary Matrix," to be published in Jour. Assn. Comp. Mach.

Berge, Claude. Théorie des Graphes et ses Applications. Paris, Dunod, 1958. Translated into English by Alison Doig, The Theory of Graphs and its Applications. New York: John Wiley & Sons; Inc., 1962.

Bjerhammer, A. Trans. Royal Inst. Tech. Stockholm. No.49, 1951, pp. 1-86.

Bonini, Charless P. "Accounting-Information Systems in the Firm." ONR Research Memorandum, No. 59, Carnegie Institute of Technology, December, 1958.

______. "Simulation of Information and Decision Systems in the Firm." ONR Research Memorandum No. 77, Carnegie Institute of Technology, January, 1961.

______. Simulation of Information and Decision. Ph.D. Thesis. Pittsburgh: Carnegie Institute of Technology, May, 1962.

Bott, R. and R.J. Duffin. "On the Algebra of Networks," Transactions of the American Mathematical Society, LXXIV, No.1, January, 1953, pp. 99-100.

Bray, F. Sewell. Four Essays in Accounting Theory. London: Oxford University Press, 1953.

Brummet, R.L. Overhead Costing. Ann Arbor: University of Michigan, 1958.

Burchard, Joseph R. "Critical Look at the Marginal Graph Technique," NAA Bulletin, May, 1961.

Cayley, Arthur. The Collected Mathematical Papers of Arthur Cayley. Cambridge: University Press, 1895. Vols. I-XIII plus index volume.

______. The Principles of Book-keeping by Double Entry. Cambridge: at the University Press, 1907.

Chambers, R.J. Towards a General Theory of Accounting. Sydney: The University of Sydney, 1961.

Charnes, A. "Optimality and Degeneracy in Linear Programming," Econometrica, XX, No.2, pp. 160-170, (1952).

______ and W.W. Cooper. "Deterministic Equivalents for Different Objectives in Chance Constrained Programming," ONR Research Memorandum No. 37. The Technological Institute, Northwestern University, December, 1960.

______ and ______. Management Models and Industrial Applications of Linear Programming. New York: John Wiley and Sons, Inc., 1961.

______ and ______. "Making Models Mathematical," a Mimeograph. Pittsburgh, Carnegie Institute of Technology, November, 1962.

_______ and _______. "Constrained Extremization Models and Their Use in Developing System Measures," a Mimeograph. Pittsburgh: Carnegie Institute of Technology, April, 1963.

_______, _______, and A. Henderson. An Introduction to Linear Programming. New York: John Wiley & Sons, Inc., 1953.

_______, _______, and Y. Ijiri. "Breakeven Budgeting and Programming to Goals," Journal of Accounting Research, I, No.1, May, 1962.

_______, _______, and M.H Miller. "Application of Linear Programming to Financeal Budgeting and the Costing of Funds," Journal of Business of the University of Chicago, XXXII, No.1, January, 1959.

_______, _______, and G.L. Thompson. "Constrained Generalized Medians and Linear Programming under Uncertainty," ONR Research Memorandum No.84. Pittsburgh; Carnegie Institute of Technology, July 6, 1961.

_______, _______, and _______. "Critical Path Analyses via Chance Constrained and Stochastic Programming," ONR Research Memorandum No.101. Pittsburgh: Carnegie Institute of Technology, July, 1962.

_______, _______, and _______. "Chance Constrained Studies for Linear Programming Under Uncertainty; Part One; Constrained Generalized Medians and Two-Stage Problems with General Linear Structure: Part Two; Linear Decision Rules and General K-Stage Problems," ONR Research Memorandum No.106. Pittsburgh: Carnegie Institute of Technology, December, 1962.

Churchill, N.C. Behavioral Effects of an Audit: An Experimental Study. Ph.D. dissertation. Ann Arbor: University Microfilms, 1962.

_______ and W.W. Cooper. "Effects of Auditing Records: Individual Task Accomplishment and Organization Objectives," ONR Research Memorandum No.98. Pittsburgh; Carnegie Institute of Technology, June, 1962.

Cline, R.E. "On the Computation of the Generalized Inverse A^+ of an Arbitrary Matrix A, and the Use of Certain Associated Eigenvectors in Solving the Allocation Problem: Lafayette, Ind.: Purdue University, 1958.

_______. "Generalized Inverse Representation for Matrices Partitioned as A = [U, V]." The University of Michigan, 1962.

Cohen, K.J., et. al., "The Carnegie Tech Management Game," The Journal of Business of the University of Chicago, XXXIII, No.4, Oct., 1960.

Cooper, W.W. Historical and Alternative Costs: A Study in Some Relations between the Economic Theory of the Firm and Accounting Control of Operations. Pittsburgh: Carnegie Institute of Technology, 1950.

_______. "Some Implications of the Newer Analytic Approaches to Management," California Management Review, IV, Fall, 1961.

Crivelli, Pietro. An Original Translation of the Treatise on Double-Entry Book-keeping by Frater Lucas Pacioli, Published by The Institute of Book-keepers, Ltd; 133, Moorgate, London, E.C.2, 1924.

Cyert, R.M., Justin Davidson and G.L. Thompson. "Estimation of the Allowance for Doubtful Accounts by Markov Chains." Management Science, April, 1962.

_______ and James G. March. A Behavioral Theory of the Firm. Englwood Cliffs, N.J.: Prentice-Hall, Inc., 1963.

Dantzig, G.B. and P. Wolfe. "Decomposition Principle for Linear Programs," Operations Research VIII, No.1, February, 1960.

Davidson, Sidney. "Old Wine Into New Bottles," Accounting Review, April, 1963, pp. 278-84.

Dean, Joel. "Cost Structure of Enterprises and Break-Even Charts," American Economic Review Supplement, 1947.

_______. "Methods and Potentialities of Break-Even Analysis," The Australian Accountant XXI, Sos. 10 and 11, October and November, 1951.

_______. "Managerial Economics," in William Grant Ireason and Eugene L. Grant, eds., Handbook of Industrial Engineering Management. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1955.

den Broader, G.G. and A. Charnes. "Contributions to the Theory of a Generalized Inverse for Matrices," Research Memorandum, ONR Research Project on Planning and Control of Industrial Operations. Lafayetts: Purdue University, February, 1957.

Devine, Carl T. "Boundaries and Potentials of Reporting on Profit-Volume Relationships," NAA Bulletin, Sec. 1, January, 1961, pp. 5-14.

Dill, W.R., et. al. "Experiences with a Complex Management Game," California Management Review, III, No.3, Spring, 1961, pp. 39-51.

Dorfman, R. Application of Linear Programming to the Theory of the Firm. Berkeley: University of California Press, 1951.

_______, Paul A. Samuelson, and Robert M. Solow. Linear Programming and Economic Analysis. New York: McGraw-Hill Book Co., Inc., 1958.

Duffin, R.J. Archive for Rat. Mech. and Anal. IX, No.4, 1962 pp. 309-18.

Eastwood, R. Parker. "The Break-Even Chart as a Tool for Managerial Control" in Practical Uses of Break-Even and Budget Controls. Production Series No.186. New York: American Management Association, 1949.

Eckholdt, John L. "Using the Break Even Chart in Product-Line Decisions," NAA Bulletin, July, 1960, pp. 43-50.

Edwards, R.S. A Survey of the French Contributions to the Study of Cost Accounting During the 19th Century. London: Gee and Co., 1937.

Gardner, Fred V. Profit Management and Control. New York: McGraw-Hill Book Co., Inc., 1955. Especially Section II, Chapters 12-27.

Garner, S. Paul. "Historical Development of Cost Accounting," The Accounting Review, XXII, 4, October, 1947, pp. 385-389.

_______. "Has Cost Accounting Come of Age? NAA Bulletin XXXIII, No.3, November, 1951, pp. 387-392.

_______. Evolution of Cost Accounting to 1925. Alabama: University of Alabama Press, 1954.

Gassner, L.M, E.B. Magee and E.R. Maurer. "Linear Programming as a Method for the Analysis of the Break-Even Point of Sales volume." Ditto, Term project for the course, Advanced Business and Engineering Economics, at the Graduate School of Industrial Administration. Pittsburgh: Carnegie Institute of Technology, May, 1955.

Geijsbeek-Molenaar, L.B. Ancient Double Entry Bookkeeping. Denver, Colorado, 1914.

Gould, J.R., "The Economist's Cost Concept and Business Problems," in W.T. Baxter and S. Davidson (eds.)., Studies in Accounting Theory, 2nd ed. Homewood, Illinois: Richard D. Irwin, Inc., 1962.

Guillemin, Ernest A. Introductory Circuit Theory. New York: John Wiley and Sons, Inc., 1953.

Halmos, Paul R. Finite-Dimensional Vecotr Spaces. Second Ed. Princeton, N.J.: D. Van Nostrand Co., Inc., 1958.

Hammer, Frederick S. The Demand for Physical Capital: Application of a Wealth Model. Ph.D. Thesis. Pittsburgh: Carnegie Institute of Technology, 1963.

Hatfield, Henry Rand. "An Historical Defense of Book-keeping," Journal of Accountancy, 37 (1924) reprinted in W.T. Baxter, ed., Studies in Accounting. London: 1950

Hill, Thomas M. and Myron J. Gordon. Accounting: A Management Approach. Homewood, Illinois: Richard D. Irwin, Inc., 1959.

Hitch, C.J. "Sub-Optimization Problems," Journal of the Operations Research Society, I, No.2, May, 1953, pp. 87-99.

Ijiri, Yuji. An Application of Input-Output Analysis to Some Problems in Cost Accounting, Unpublished Term Paper. Minneapolis: University of Minesota, 1960.

_______, F.K. Levy, and R.C. Lyon. "A Linear Programming Model for Budgeting and Financial Planning," ONR Research Memorandum No.94. Pittsburgh: Carnegie Insitute of Technology, July, 1962. To be published in Journal of Accounting Research, I.

Jackson, J.G.C. "The History of Methods of Exposition of Double-entry Book-keeping in England" in A.C. Littleton and B.S. Yamey, (eds.) Studies in the History of Accounting. Homewood, Illinois: Richard D. Irwin, Inc., 1956, pp. 288-312.

Jaedicke, Robert K. "Improving Break-Even Analysis by Linear Programming Techniques," NAA Bulletin. Sec. 1, March, 1961, pp. 5-12.

_______. "Some Notes on Product-Combination Decisions," The Accounting Review, XXXIII, No.4, pp. 586-601.

Kemeny, J.G., Hazleton Mirkil, J.Laurise Snell, and Gerald L. Thompson. Finite Mathematical Structures. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1959.

_______, Arthur Schleifer, Jr., J. Laurie Snell, Gerald L. Thompson. Finite Mathematics with Business Applications. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962.

Kiene, R.C. "Mathematical Formalisms of Command, Control, and Communication Systems." Minneapolis, Minnesota: Minneapolis-Honeywell Regulator Company, Military Products Group, October 25, 1961.

Kohler, Eric L. A Dictionary for Accountants. Englewood Cliffs, N.J., Prentice-Hall, Inc., 1952.

Koopmans, T.C., ed. Activity Analysis of Production and Allocation. New York: John Wiley and Sons, Inc., 1951.

Krech, David and Richard S. Crutchfield. Theory and Problems of Social Psychology. New York: McGraw-Hill Book Co., Inc., 1948.

Kuhn, H. and A. Tucker. "Nonlinear Programming," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Ed., J. Neyman. Berkeley: University of California Press, 1951.

Lang, Theodore (ed.). Cost Accountants' Handbook. New York: The Ronald Press Company, 1952.

Leontief, Wassily W. The Structure of American Economy, 1919 - 1939: an Experimental Application of Equilibrium Analysis, (2nd. ed.). New York, Oxford University Press, 1951.

Lerner, Abba P. The Economics of Control: Principles of Welfare Economics. New York: The MacMillan Company, 1946.

Littleton, A.C. Accounting Evolution to 1900. New York: American Institute Publishing Co., 1933.

_______ and B.S. Yamey, eds. Studies in the History of Accounting. Homewood, Ill., Richard D. Irwin, Inc., 1956.

Machol, R.E., ed. Information and Decision Processes. New York: McGraw-Hill Book Co., Inc., 1960. Especially, Rosenblatt, D., "On Some Aspects of Models of Complex Behavioral Systems," and Rosenblatt, M., "An Aggregation Problem for Markov Chains."

March, J.G. and H.A. Simon. Organizations. New York: John Wiley & Sons, Inc., 1958.

Marschak, Jacob. "Problems in Information Economics," Working Paper No. 24, Western Management Science Institute, University of California, Los Angeles, November 25, 1962.

_______, "Towards an Economic Theory of Organization and Information," in R.M. Thrall, et. al. (eds.), Decision Processes.

Mattessich, Richard. "Towards a General and Axiomatic Foundation of Accountancy," Accounting Research, October, 1957.

______. "Budgeting Models and System Simulation," Accounting Review, July, 1961, pp. 384-97.

Mautz, R.K. "Accounting as a Social Science," Accounting Review, XXXVIII, No.2, April, 1963, pp. 317-325.

May, P.A. "Profit Polygraph for Product Mix Analysis," NAA Bulletin, November, 1955.

Mellon, W.G. An Approach to a General Theory of Priorities. Princeton, N.J.: Princeton University, July, 1962.

Moonitz, Maurice. The Basic Postulates of Accounting, An Accounting Research Study No.1. New York: American Institute of Certified Public Accountants, 1961.

Moore, E.H. Bull. Amer. Math. Soc. XXVI, 1920, pp. 394-395.

Motzkin, T.S., H. Raiffa, G.L. Thompson, and R.M. Thrall. "The Double Description Method," in Contributions to the Theory of Games, Vol. II (Editor, H.W. Kuhn and A.W. Tucker). Princeton, N.J.: Princeton University Press, 1953, pp. 51-73.

Murphy, Gardner. "Social Motivation," Handbook of Social Psychology, Gardner Lindzey, ed. Cambridge: Addison-Wesley Publishing Co., Inc., 1954.

Orden, Alex. "Minimization of Indefinite Quadratic Functions with Linear Constraints," a Paper presented at a Symposium on Mathematical Programming, held at the University of Chicago, June 18-22, 1962.

Paton, W.A. Recent and Prospective Developments in Accounting Theory. (Business Research Studies No.25) Boston: Harvard University, Graduate School of Business Administration, March, 1940.

______ and A.C. Littleton. An Introduction to Corporate Accounting Standards. New York: American Accounting Association, 1940.

Patrick, A.W. "Some Observations on the Break-Even Chart," The Accounting Review 33, No.4, pp. 573-580.

Penrose, R. "A Generalized Inverse for Matrices," Proceedings of the Cambridge Philosophical Society, LI, 1955, pp. 406-13.

Peragello, Edward. Origin and Evolution of Double Entry Accounting. New York: American Institute Publishing Co., 1938.

Pigman, Nathaniel M., Jr. "Simplified Financial Research--An Example in Profit Maximization," NAA Bulletin, Sec. 2, January, 1962, pp. 87-92.

Puzey, R.V. "Accounting is Communication," Journal of Accountancy, September, 1961.

Pye, M.L. "How to Determine Break-Even Points with simple Algebraic Formulas," Journal of Accountancy, August, 1948, pp. 133-137.

Richards, Allen B. "Input-Output Accounting for Business," The Accounting Review, XXXV, July, 1960, pp. 429-436.

Robertson, H.M. Aspects of the Rise of Economic Individualism. Cambridge: University Press, 1935.

Rosenblatt, D. "On Aggregation and Consolidation in Finite Substochastic Systems, I, II, III, IV," Abstracts Annals of Mathematical Statistics, XXVIII, 1957, pp. 1060-61.

Rosenblatt, D. "On Aggregation and Consolidation in Linear Systems." Technical Report C, Office of Naval Research Contract Nonr-1180(00) NR-047-012. Washington, The American University Department of Statistics, August, 1956.

Samuelson, Paul A. Foundations of Economic Analysis. Cambridge: Harvard University Press, 1947.

Sanders, Thomas H., Henry R. Hatfield, and Underhill Moore. A Statement of Accounting Principles. New York: American Institute of Accountants, 1938.

Schlatter, Charles F. Advanced Cost Accounting. New York: John Wiley and Sons, Inc., 1939.

Shillinglaw, Gordon. "Problems in Divisional Profit Measurement," NAA Bulletin, Sec. 1, March, 1961, pp. 34-43.

Simon, Herbert A. Administrative Behavior. Second Edition. New York: The Macmillan Co., 1959.

_______ and Albert Ando. "Aggregation of Variables in Dynamic Systems," Econometrica, XXIX, April, 1961, pp. 111-38.

Solomon, Ezra. (ed.). The Management of Corporate Capital. Chicago; Graduate School of Business, The University of Chicago, 1959.

Solomons, David, ed. Studies in Costing. London: Sweet and Maxwell, Ltd., 1952.

Sombart, Werner. Der Moderne Kapitalismus. Munchen und Leipzig: von Duncker und Humblot, 1922.

Spencer, Milton H. "Axiomatic Method and Accounting Science," Accounting Review, April, 1963, pp. 310-316.

Sprouse, Robert T. and Maurice Moonitz. A Tentative Set of Broad Accounting Principles for Business Enterprises, an Accounting Research Study No. 3. New York: American Institute of Certified Public Accountants, 1962.

Stedry, A. Budget and Control and Cost Behavior. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1960.

_______ and A. Charnes. "Exploratory Models in the Theory of Budget Control." ONR Research Memorandum No. 43. Evanston: The Technological Institute, Northwestern University, January, 1962.

Steinman, L. Data Processing: The Extension of Man's Sensors and Physical Capabilities (AMELCO Inc., Report No.2). Los Angels, AMELCO Inc., September, 1962.

Theil, H. "Alternative Approaches to the Aggregation Problem," Reprint Series No. 57, Logic, Methodology and Philosophy of Science, Netherlands School of Economics, Econometric Institute.

Thirlby, G.E. "Notes on the Maximization Process in Company Administration, Economica, XVII, 1950, 266-282.

Trueblood, Robert M. "Internal Control," a Mimeographed Internal Memorandum. Pittsburgh: Touche, Niven, Bailey and Smart, 1950.

_______ and H. Justin Davidson. "Accounting for Decision-Making," Accounting Review, XXXVI, October, 1961, pp. 577-82.

Tse, John Y.D. Profit Planning through Volume Cost Analysis. New York: The MacMillan Co., 1960.

Tucker, A.W. Condensed Schemata for Dantzig's Simplex Mehtod, Mimeorgraphed. Princeton University, January, 1958.

Vancamp, H.K. "The Use of the Product Profit and Loss Budget for Marginal Sales Decision," NAA Bulletin, Sec. 1. November, 1960, pp. 5-16.

Vickers, Douglas. "On the Economics of Break-Even 'Input Output'" The Accounting Review, XXXV, No.3, July, 1960, pp. 405-412.

Wanger, Harvey M. Statistical Management of Inventory Systems. New York: John Wiley & Sons, 1962.

Wengartner, H.M. Mathematical Programming and the Analysis of Capital Budgeting Problems. Ph.D. Thesis. Pittsburgh: Carnegie Institute of Technology, 1961.

Whinston. A. Price Coordination in Decentralized Systems. Ph.D Thesis. Pittsburgh: Carnegie Institute of Technology, June, 1962.

Willson, James D. "Practical Applications of Cost-Volume-Profit Analysis," NAA Bulletin, Sec. 1, No.7, March, 1960, pp. 5-18.

Zannetos, Zenon S. "Mathematics as a Tool of Accounting Instruction and Research." Accounting Review, April, 1963, pp.326-335

Zelen, M. "The Role of Constraints in the Theory of Least Squares," MRC Report No.314, Mathematics Research Center, U.S. Army, The University of Wisconsin, Madison, Wis., 1962.

UNCLASSIFIED

UNCLASSIFIED